

STIC-Biotech/ChemLib

146244

From: Slobodyansky, Elizabeth
Sent: Friday, February 25, 2005 9:40 PM
To: STIC-Biotech/ChemLib
Subject: 10/019,284 SEQ

Please search for case 10/019,284:

ES

SEQ ID NOs: 1 and 2 against commercail ~~and interference~~ databases.

Thank you.

Elizabeth Slobodyansky, PhD

Primary Examiner
Art Unit 1652
REM 2D83
571-272-0941
MAILBOX 2C70

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: 3/1/05
Date Completed: 3/1/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA Sequence: # _____
AA Sequence: # 1 _____
Structure: # _____
Bibliographic: _____
Litigation: _____
Patent Family: _____
Other: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: depts
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 13:51:53 ; Search time 24423 Seconds

(without alignments)
11842.492 Million cell updates/sec

Title: US-10-019-284B-1

Perfect score: 5969

Sequence: 1 agtcgcgtcagccaccatt.....tgatatttcgcgcctgaa 5969

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: *
1: gb_ba: *
2: gb_ncg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pac: *
7: gb_ph: *
8: gb_dl: *
9: gb_pr: *
10: gb_ro: *
11: gb_stb: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5969	100.0	5969	6	BD093238
2	5556.2	94.8	320550	1	AP005282
3	5556.2	94.8	349136	1	EX927155
4	5556.2	94.8	349980	6	AX127152
5	3687	61.8	6911	6	E11760
6	3687	61.8	6911	6	I26124
7	1874.2	31.4	1983	6	BD165105
8	1874.2	31.4	1983	6	AX122988
9	1457	24.4	1656	6	BD094213
10	1412.6	23.7	1527	6	AX069134
11	1412.6	23.7	1527	6	AX469820
12	1321.2	22.1	1342	6	AX065221
13	1267.8	21.2	1299	6	BD165106
14	1267.8	21.2	1299	6	AX122989
15	1237.4	20.7	1287	6	AX065287
16	1237.4	20.7	1287	6	AX065289
17	1110.4	18.6	1152	6	BD165108
18	1110.4	18.6	1152	6	AX122991
19	1049.8	17.6	1109	6	AX069136

20	762	12.8	882	6	AX065297
21	671.8	11.3	759	6	BD165107
22	671.8	11.3	759	6	AX122990
23	474.2	7.9	479	6	BD094149
24	466.6	7.8	498	6	BD094148
25	433	7.3	343050	1	AL935252
26	421.6	7.1	21838	1	PDCAPOPER
27	421.6	7.1	21839	1	PPSUFOP
28	397	6.7	5800	1	LL29701S
29	371.8	6.2	341553	1	BX248355
30	353.2	5.9	2508	1	STRSCRA
31	352.8	5.9	300330	1	AP005222
32	351.6	5.9	9979	1	AE005011
33	343	5.7	4750	1	AY177419
34	326.6	5.5	110000	1	AE016827_07
35	325	5.4	2655	1	SXSCRA
36	323	5.4	110000	1	BX571856_25
37	322.8	5.4	177911	1	AE017206
38	322.8	5.4	349980	6	AX926712
39	320.2	5.4	10264	1	AE004395
40	320	5.4	2788	1	AF401046
41	319.8	5.4	110000	1	BX571857_24
42	319.8	5.4	301250	1	AP004830
43	318.2	5.3	297850	1	AP003137
44	318.2	5.3	342350	1	AP003365
45	316.6	5.3	39425	1	AC090968

ALIGNMENTS

RESULT 1	BD093238	5969 bp	DNA	linear	PAT 27-AUG-2002
LOCUS	BD093238				
DEFINITION	DNA encoding sucrose PTS enzyme II.				
ACCESSION	BD093238				
VERSION	BD093238.1 GI:22638826				
KEYWORDS	WO 0102584-A/1.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 5969)				
AUTHORS	Izui, M., Sugimoto, M., Nakamatsu, T. and Kurahashi, O.				
TITLE	DNA encoding sucrose PTS enzyme II				
JOURNAL	Patent: WO 0102584-A 1 11-JAN-2001.				
COMMENT	OSAMU KURAHASHI				
	OS Brevibacterium lactofermentum				
	PV WO 0102584-A/1				
	PD 11-JAN-2001				
	PF 30-JUN-2000 WO 2000JP004348				
	PR 02-JUL-1999 JP 99P 189512				
	PI MASAKO IZUI, MASAKAZU SUGIMOTO, TSUYOSHI NAKAMATSU, OSAMU KURAHASHI				
	PC C12N15/54, C12N9/12				
	CC				
	EH				
	FT				
FEATURES	source	1..5969			
		Location/Qualifiers			
		Location/Qualifiers			
		Key			
		Key			
		CDS			
ORIGIN		1..5969			
		/organism="unidentified"			
		/mol_type="genomic DNA"			
		/db_xref="taxon:32644"			
Query Match	100.0%	Score 5969	DB 6	Length 5969	
Best Local Similarity	100.0%	Pred. No. 0			
Matches 5969	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
QY	1	AGTCCGTCGACGCGCACATTGATGTGTGTGTCACGAGCTTGGAGGCTTTTACATCT	60		
Db	1	AGTCCGTCGACGCGCACATTGATGTGTGTGTCACGAGCTTGGAGGCTTTTACATCT	60		

QY 61 ACGCTCCGTCGCGAGTGGAGTCAATTACGGGTGGATCAAGCCGCTGAAGTTGCG 120
 Db 61 ACGCTCCGTCGCGAGTGGAGTCAATTACGGGTGGATCAAGCCGCTGAAGTTGCG 120
 QY 121 GAACCCATGTTCTCTTGTGGTTGAGGGAACAGTGGGGTGAAGTTTTCGAAG 180
 Db 121 GAACCCATGTTCTCTTGTGGTTGAGGGAACAGTGGGGTGAAGTTTTCGAAG 180
 QY 181 TCTGAGTTTAAAGTTATGATCATCAGCTTGAAGGCTGAGTAACTCAGTAGACCTG 240
 Db 181 TCTGAGTTTAAAGTTATGATCATCAGCTTGAAGGCTGAGTAACTCAGTAGACCTG 240
 QY 241 CAACAGCAGGCTCAAGTCCGAAGATAATTAACTTAGATCCGTAGACATTAACATCAT 300
 Db 241 CAACAGCAGGCTCAAGTCCGAAGATAATTAACTTAGATCCGTAGACATTAACATCAT 300
 QY 301 CGTCCATGCTTGGTGAAGGAACCAATACCTGAGAAAGATGGAGAAAGTGTGCAAT 360
 Db 301 CGTCCATGCTTGGTGAAGGAACCAATACCTGAGAAAGATGGAGAAAGTGTGCAAT 360
 QY 361 ATCAAGAAATGAGGCTCAAGCATTAATAAATTAGAGGAAGATTGTCCTCCCTCG 420
 Db 361 ATCAAGAAATGAGGCTCAAGCATTAATAAATTAGAGGAAGATTGTCCTCCCTCG 420
 QY 421 GGGTGAATGATGCTTCTTCACTGAAACGGCATCATACGGAACCTCTGAGAAC 480
 Db 421 GGGTGAATGATGCTTCTTCACTGAAACGGCATCATACGGAACCTCTGAGAAC 480
 QY 481 CAGACCTTAAACGAGAGATTCACCCCGAACTCCCAAGATGTTTCCGGTTTAAATG 540
 Db 481 CAGACCTTAAACGAGAGATTCACCCCGAACTCCCAAGATGTTTCCGGTTTAAATG 540
 QY 541 ATCTTATATCAAGGTGGAACGGTGGCGGTTTCTTAAGGAAACGACAGACAGGCGA 600
 Db 541 ATCTTATATCAAGGTGGAACGGTGGCGGTTTCTTAAGGAAACGACAGACAGGCGA 600
 QY 601 GGAACACCGGAGTATCACCCGGAACATGGCAACGCGATGTTGCCAAGCATGTTT 660
 Db 601 GGAACACCGGAGTATCACCCGGAACATGGCAACGCGATGTTGCCAAGCATGTTT 660
 QY 661 CGGCGCGGCTGACGCACTGGACGCGAGGTGGAACCTTATTTCCCTGTGTGAAGAG 720
 Db 661 CGGCGCGGCTGACGCACTGGACGCGAGGTGGAACCTTATTTCCCTGTGTGAAGAG 720
 QY 721 TCTGCTGTGGGATTCACCTGAGAGGCCCTTTCATCAACGATGCCGTTGTGTGCTC 780
 Db 721 TCTGCTGTGGGATTCACCTGAGAGGCCCTTTCATCAACGATGCCGTTGTGTGCTC 780
 QY 781 AAAACCGGATTCATTTTCCGGGAACCCAAACAGATCTTGCCTGGGTGATCATGCG 840
 Db 781 AAAACCGGATTCATTTTCCGGGAACCCAAACAGATCTTGCCTGGGTGATCATGCG 840
 QY 841 GAAAAGTTGATCAATTCATCAAGTAGCGCGGAACTGACATCTTGTGAGCTTC 900
 Db 841 GAAAAGTTGATCAATTCATCAAGTAGCGCGGAACTGACATCTTGTGAGCTTC 900
 QY 901 TCGATCTCTGCGAGCGGCAACATCATTTGCTTCTTCCGGCAACATGAGCAATTTG 960
 Db 901 TCGATCTCTGCGAGCGGCAACATCATTTGCTTCTTCCGGCAACATGAGCAATTTG 960
 QY 961 ATACCACTACAGGCGCAATGCTTGGCTTAAAGAAAATGTGACGGTCAACGGCTAC 1020
 Db 961 ATACCACTACAGGCGCAATGCTTGGCTTAAAGAAAATGTGACGGTCAACGGCTAC 1020
 QY 1021 ATTGTTCAATGCGATGCTTCGCTGATCATAGGCTCCCGGAGGTGAGGCTTTC 1080
 Db 1021 ATTGTTCAATGCGATGCTTCGCTGATCATAGGCTCCCGGAGGTGAGGCTTTC 1080
 QY 1081 TTGCTGGGCAAGTCCGGGGAACGATATGTTGATGATGCGGAGGCTGATTTG 1140
 Db 1081 TTGCTGGGCAAGTCCGGGGAACGATATGTTGATGATGCGGAGGCTGATTTG 1140
 QY 1141 CCGATGGAACGGTGCATGCTAGCTGTTCAACAAAGCCTTTTTCATCAAGACCGCATG 1200

Db 1141 CCGATGGAACGGTGCATGCTAGCTGTTCAACAAAGCCTTTTTCATCAAGACCGCATG 1200
 QY 1201 AAGCGCGGAAATGCAAGCGGTAGTAACTTTTGGGCTTTTGAACGTACCGTCAACG 1260
 Db 1201 AAGCGCGGAAATGCAAGCGGTAGTAACTTTTGGGCTTTTGAACGTACCGTCAACG 1260
 QY 1261 ATGAGTTCGCGGCTGCGGATGAGCGGCGCATTCGCGGGGACAGCACTAGCGA 1320
 Db 1261 ATGAGTTCGCGGCTGCGGATGAGCGGCGCATTCGCGGGGACAGCACTAGCGA 1320
 QY 1321 GTCACTGTCGACCAAGTGGAGGATGACGCTTATGACGCGACCTTCACACT 1380
 Db 1321 GTCACTGTCGACCAAGTGGAGGATGACGCTTATGACGCGACCTTCACACT 1380
 QY 1381 CAACCGTCCGCTTAAATCTCGGCTTGGCCATCAAGAAATGCTAAATCCAACTCG 1440
 Db 1381 CAACCGTCCGCTTAAATCTCGGCTTGGCCATCAAGAAATGCTAAATCCAACTCG 1440
 QY 1441 CAATTTTGTGCTTTGACTCAACCGGCGAGGTGCAAGAGTCACTTAACTATCAAG 1500
 Db 1441 CAATTTTGTGCTTTGACTCAACCGGCGAGGTGCAAGAGTCACTTAACTATCAAG 1500
 QY 1501 TACTTAAAGTACGATAACTATCTGATTTTAAAGAGTCCCAATGGAATCACTA 1560
 Db 1501 TACTTAAAGTACGATAACTATCTGATTTTAAAGAGTCCCAATGGAATCACTA 1560
 QY 1561 TCTGCAAGACGAGCAGAAATGCGCAACGATGTCATGCTTAATGCAACCTTCGCA 1620
 Db 1561 TCTGCAAGACGAGCAGAAATGCGCAACGATGTCATGCTTAATGCAACCTTCGCA 1620
 QY 1621 ACAAGGATGGAACCTTGGGCTTGAACAAGATCCTCAACGATGATCTTAAACCAAG 1680
 Db 1621 ACAAGGATGGAACCTTGGGCTTGAACAAGATCCTCAACGATGATCTTAAACCAAG 1680
 QY 1681 TCATTGCGATGATGAACTGAGGAGTGTCAATTCAGAACTGCAAGGCAATTTGTTGG 1740
 Db 1681 TCATTGCGATGATGAACTGAGGAGTGTCAATTCAGAACTGCAAGGCAATTTGTTGG 1740
 QY 1741 ATGAATAGTGGAGCTAACCCGTCAGATGAAACAGCTTCTTAAACCAATCGCAAG 1800
 Db 1741 ATGAATAGTGGAGCTAACCCGTCAGATGAAACAGCTTCTTAAACCAATCGCAAG 1800
 QY 1801 AGTTCACTGACCAATGCAATCTGTTGATGAAAGGTCTACAGCCAGATGGTGAAC 1860
 Db 1801 AGTTCACTGACCAATGCAATCTGTTGATGAAAGGTCTACAGCCAGATGGTGAAC 1860
 QY 1861 CTGATTCATGAGAGCAGTGTGAGATGAGGCAAAAGTGGTGCAGAAATCGTTGAAG 1920
 Db 1861 CTGATTCATGAGAGCAGTGTGAGATGAGGCAAAAGTGGTGCAGAAATCGTTGAAG 1920
 QY 1921 TTCAATCCTTGGCATGCGCGGAAACGCGCAATCGCTTTCATTTGAACATATCTTCT 1980
 Db 1921 TTCAATCCTTGGCATGCGCGGAAACGCGCAATCGCTTTCATTTGAACATATCTTCT 1980
 QY 1981 GTCAAGACTGACAAAGGTTCAGGCGCTGCAACCTTAACTGTGAGGACAAAGCTGAT 2040
 Db 1981 GTCAAGACTGACAAAGGTTCAGGCGCTGCAACCTTAACTGTGAGGACAAAGCTGAT 2040
 QY 2041 CTTCAACACATGCAAGAGTCCCAACCGCGCTCACCGGATTTGGGCACTTTGTC 2100
 Db 2041 CTTCAACACATGCAAGAGTCCCAACCGCGCTCACCGGATTTGGGCACTTTGTC 2100
 QY 2101 CCGCGGCAAAACATGCTGTGTGAGCACTGTGAAAGAAAAGCGCATCCGCG 2160
 Db 2101 CCGCGGCAAAACATGCTGTGTGAGCACTGTGAAAGAAAAGCGCATCCGCG 2160
 QY 2161 AACTGTGAAGGCCAGTGACTGCTTCCAGAGTTTCACTCTGTGATGCAACAT 2220
 Db 2161 AACTGTGAAGGCCAGTGACTGCTTCCAGAGTTTCACTCTGTGATGCAACAT 2220
 QY 2221 GCCACATATGTTGATGAAAGCAGTATCAAGCTGGAACGCTGATCACTACG 2280

Db 2221 GCCACATCATGTTGGATGAGACAGATATCCAGCTGGAAAAGCGTGATCACTACACG 2280
Qy 2281 TCTCATGAGCAATTAAAGCTGCGTGAAGAACAAAAGAAAGTACTGTGTGGGCTATG 2340
Db 2281 TCTCATGAGCAATTAAAGCTGCGTGAAGAACAAAAGAAAGTACTGTGTGGGCTATG 2340
Qy 2341 CACACGAACTTCCAGTTTGGGCGCTGGGTACCATGTACCTCTCCGAGAGGCAAGCTC 2400
Db 2341 CACACGAACTTCCAGTTTGGGCGCTGGGTACCATGTACCTCTCCGAGAGGCAAGCTC 2400
Qy 2401 AATGATCCCAAGCAATGTACGTGATGAGATACCTCCAGCTACTACTACAGACAGAT 2460
Db 2401 AATGATCCCAAGCAATGTACGTGATGAGATACCTCCAGCTACTACTACAGACAGAT 2460
Qy 2461 CCAAGTTTCCCTTCCGACCAAGGCGACCGGCTGGGCTTCAACAACAGCCGTTGACC 2520
Db 2461 CCAAGTTTCCCTTCCGACCAAGGCGACCGGCTGGGCTTCAACAACAGCCGTTGACC 2520
Qy 2521 GGAACGCAAGCAATGAGTGAAGGACGACCTGCCGACGCTTTTACCGGATGATCCTAT 2580
Db 2521 GGAACGCAAGCAATGAGTGAAGGACGACCTGCCGACGCTTTTACCGGATGATCCTAT 2580
Qy 2581 GACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640
Db 2581 GACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640
Qy 2641 TACACCGGCAACCTTAAATAATGACGGAAGCGCGGCAACCAAAACCTTGTGAAATC 2700
Db 2641 TACACCGGCAACCTTAAATAATGACGGAAGCGCGGCAACCAAAACCTTGTGAAATC 2700
Qy 2701 GAGGACCCCAATGAGTGAAGGACGATGATGATGATGATGATGATGATGATGATGAT 2760
Db 2701 GAGGACCCCAATGAGTGAAGGACGATGATGATGATGATGATGATGATGATGATGAT 2760
Qy 2761 GACGACCCGCGACGAGTTCACACCCATTAACGCGATCCCATGATCAGCCCTGATGAT 2820
Db 2761 GACGACCCGCGACGAGTTCACACCCATTAACGCGATCCCATGATCAGCCCTGATGAT 2820
Qy 2821 GATGTTGGAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2880
Db 2821 GATGTTGGAACAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2880
Qy 2881 TACCGTCCGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
Db 2881 TACCGTCCGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940
Qy 2941 GATGACCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
Db 2941 GATGACCAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3000
Qy 3001 AACCTTTTAAAGCTTCCGATGAGAAAGTGGGAAAGTCTGACGCTGATTTTCTGT 3060
Db 3001 AACCTTTTAAAGCTTCCGATGAGAAAGTGGGAAAGTCTGACGCTGATTTTCTGT 3060
Qy 3061 CCAAGAGATGAGACGGAATCCAGATGAGTATCTACAGCAAGCTTCTGACAGTGC 3120
Db 3061 CCAAGAGATGAGACGGAATCCAGATGAGTATCTACAGCAAGCTTCTGACAGTGC 3120
Qy 3121 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
Db 3121 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3180
Qy 3181 CTGGAATTTCCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Db 3181 CTGGAATTTCCGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Qy 3241 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
Db 3241 CTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
Qy 3301 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
Db 3301 TGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360

Qy 3361 GAGCTCTTCTCCAGAGGAGGAGTCCGGGGATATCAGATCTGATTAAGTTCTGAACCT 3420
Db 3361 GAGCTCTTCTCCAGAGGAGGAGTCCGGGGATATCAGATCTGATTAAGTTCTGAACCT 3420
Qy 3421 GTCCGAGTGAATCCAGAGCAATATTTTCCCTGAGTGGAGTGTGTCCGTTGTCTGTG 3480
Db 3421 GTCCGAGTGAATCCAGAGCAATATTTTCCCTGAGTGGAGTGTGTGTCCGTTGTCTGTG 3480
Qy 3481 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540
Db 3481 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3540
Qy 3541 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
Db 3541 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
Qy 3601 CTTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660
Db 3601 CTTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660
Qy 3661 ACAATATCTTCCCAAAATCCTTGAATCGGACCAAAATTAAGTTAATTTGTTTAC 3720
Db 3661 ACAATATCTTCCCAAAATCCTTGAATCGGACCAAAATTAAGTTAATTTGTTTAC 3720
Qy 3721 TTTTGAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3780
Db 3721 TTTTGAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3780
Qy 3781 GGCATTAAGACCTGCGGCAAGGATCTGCGGCAATGAGCGGCAATGAGCGGCAATGAT 3840
Db 3781 GGCATTAAGACCTGCGGCAAGGATCTGCGGCAATGAGCGGCAATGAGCGGCAATGAT 3840
Qy 3841 CCGCGCGGCACTGTGCAACGCGTTTACGCTGCTCAAAAGACCAAGAGATGATGAT 3900
Db 3841 CCGCGCGGCACTGTGCAACGCGTTTACGCTGCTCAAAAGACCAAGAGATGATGAT 3900
Qy 3901 TCGCCAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3960
Db 3901 TCGCCAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3960
Qy 3961 CCAAGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4020
Db 3961 CCAAGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4020
Qy 4021 CTTCAAAAGCATGCTGTGTCTCAAGAGCAAGCTCAAAAGATGTTGTGCTTAACACGCA 4080
Db 4021 CTTCAAAAGCATGCTGTGTCTCAAGAGCAAGCTCAAAAGATGTTGTGCTTAACACGCA 4080
Qy 4081 CTGCTTCAAGCTGCTGTGTGTGAAGTATGAGCGGACATTTTCTGCTCCGCTATTCAT 4140
Db 4081 CTGCTTCAAGCTGCTGTGTGTGAAGTATGAGCGGACATTTTCTGCTCCGCTATTCAT 4140
Qy 4141 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4200
Db 4141 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4200
Qy 4201 TCCGCAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4260
Db 4201 TCCGCAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4260
Qy 4261 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4320
Db 4261 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4320
Qy 4321 TTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4380
Db 4321 TTTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4380
Qy 4381 GGTAAAGGCTAGAGTGTGCGGCAATGATGATGATGATGATGATGATGATGATGATGAT 4440
Db 4381 GGTAAAGGCTAGAGTGTGCGGCAATGATGATGATGATGATGATGATGATGATGATGAT 4440

QY 4441 GTTGGTTTGAATGTTGCTCAAGCTGTGTTACAGAGGACACCGTTCCTGCTGTGTGT 4500
 DB 4441 GTTGGTTTGAATGTTGCTCAAGCTGTGTTACAGAGGACACCGTTCCTGCTGTGTGT 4500
 QY 4501 CTCTTGATTCGCGAAGCATGAGAGAGTTCTGTGACAGACGACTCATGAGCACTGACAG 4560
 DB 4501 CTCTTGATTCGCGAAGCATGAGAGAGTTCTGTGACAGACGACTCATGAGCACTGACAG 4560
 QY 4561 CTCTTGATTCAGCCAGATGTTGATCTGTGCTCAACCGGCTCTTACGTTTATGTCAT 4620
 DB 4561 CTCTTGATTCAGCCAGATGTTGATCTGTGCTCAACCGGCTCTTACGTTTATGTCAT 4620
 QY 4621 TGGTCAGCAATGCGCTGGGTGGGTGATCTGTGCGACACGCTGCGACAGGACTCTATGA 4680
 DB 4621 TGGTCAGCAATGCGCTGGGTGGGTGATCTGTGCGACACGCTGCGACAGGACTCTATGA 4680
 QY 4681 TTTGCGGTGTCAGTCTGCGCGGTCTGTGTTTGGTCTGATCTAATCAACCAATGCTTATCA 4740
 DB 4681 TTTGCGGTGTCAGTCTGCGCGGTCTGTGTTTGGTCTGATCTAATCAACCAATGCTTATCA 4740
 QY 4741 TGGTCTGACCAAGTCTTCCCGCAATGAGCTGAGAGCTGTTCAACCAAGGTTGATCTT 4800
 DB 4741 TGGTCTGACCAAGTCTTCCCGCAATGAGCTGAGAGCTGTTCAACCAAGGTTGATCTT 4800
 QY 4801 CATCTTCGCAACCGCATCCATGAGCAATATCGCGAGGGTGCAGCATGTTTGGCAGTGT 4860
 DB 4801 CATCTTCGCAACCGCATCCATGAGCAATATCGCGAGGGTGCAGCATGTTTGGCAGTGT 4860
 QY 4861 CTCTCTAGGAGAGTGAAGCTCAAGGGCTTGCAGAGTGTCTTCCGCTGT 4920
 DB 4861 CTCTCTAGGAGAGTGAAGCTCAAGGGCTTGCAGAGTGTCTTCCGCTGT 4920
 QY 4921 TCTTGATTAAGAGCCCTGCGATCTTGCATGTAACCTTCCGCTGCTGCGCTTCTA 4980
 DB 4921 TCTTGATTAAGAGCCCTGCGATCTTGCATGTAACCTTCCGCTGCTGCGCTTCTA 4980
 QY 4981 CATTTGATTCGCTACCGCAGTATCGTGTGCGCTTGTGATGCACTTGTGATCAAGGC 5040
 DB 4981 CATTTGATTCGCTACCGCAGTATCGTGTGCGCTTGTGATGCACTTGTGATCAAGGC 5040
 QY 5041 AGTTGCGTTGGGCGCTGCGAGTTTCTTGGGTGTTTCTATGATGCTCAAGATATGT 5100
 DB 5041 AGTTGCGTTGGGCGCTGCGAGTTTCTTGGGTGTTTCTATGATGCTCAAGATATGT 5100
 QY 5101 CATGTTCTTGTGTTGCGCGGTGATACCTTGTGATCGATTCGCGCGAGCATTCCTTA 5160
 DB 5101 CATGTTCTTGTGTTGCGCGGTGATACCTTGTGATCGATTCGCGCGAGCATTCCTTA 5160
 QY 5161 TGGCTTTTACTTGTTCGCGCGAAGCGCAGATTCATGACAGATCGCTCCAGT 5220
 DB 5161 TGGCTTTTACTTGTTCGCGCGAAGCGCAGATTCATGACAGATCGCTCCAGT 5220
 QY 5221 GCTTCAGAGAACGACCAAGCCGAAAGCAGAGACACCCGAGATTTTCAAGATTCAC 5280
 DB 5221 GCTTCAGAGAACGACCAAGCCGAAAGCAGAGACACCCGAGATTTTCAAGATTCAC 5280
 QY 5281 CATCATCAAGGACCTTTGACCGGTGAGCTATCCCATGACAGACGTCAGAGTCCAT 5340
 DB 5281 CATCATCAAGGACCTTTGACCGGTGAGCTATCCCATGACAGACGTCAGAGTCCAT 5340
 QY 5341 GTTTCAGAGGAAAGCTTGTGCTCAGAGTTCGATCGTCCCAACAAAGGCGACGCTGT 5400
 DB 5341 GTTTCAGAGGAAAGCTTGTGCTCAGAGTTCGATCGTCCCAACAAAGGCGACGCTGT 5400
 QY 5401 TTTACAGTGAAGCGAAGATGTTGTGTGCTTCCATCTGTGATGAGCTTTGCAAGTCCG 5460
 DB 5401 TTTACAGTGAAGCGAAGATGTTGTGTGCTTCCATCTGTGATGAGCTTTGCAAGTCCG 5460
 QY 5461 CACTAAGCTGAGATGTTTCAATGTGATATCTTGAATGCACTTGTGTTGAACCGT 5520
 DB 5461 CACTAAGCTGAGATGTTTCAATGTGATATCTTGAATGCACTTGTGTTGAACCGT 5520
 QY 5521 AAACCTCAACGCGACGACTTTAACCCGCTGAAGAGCAGGCGATGAAGTCAAGCAGG 5580

DB 5521 AAACCTCAACGCGACGACTTTAACCCGCTGAAGAGCAGGCGATGAAGTCAAGCAGG 5580
 QY 5581 GAGAGCTGTGTGTGAATTCGATATTCATGATTCATTAAGCTGACGTTATGAGTAAACAC 5640
 DB 5581 GAGAGCTGTGTGTGAATTCGATATTCATGATTCATTAAGCTGACGTTATGAGTAAACAC 5640
 QY 5641 GCCGATTTGTTTGAATTCAGAGAAACCGGACCTGTAACCACTTACGTTTGGGCGA 5700
 DB 5641 GCCGATTTGTTTGAATTCAGAGAAACCGGACCTGTAACCACTTACGTTTGGGCGA 5700
 QY 5701 AATTGAACGGGAGGCAACCTGCTCAACGTCGCAAGAAAGAGCGGTGCCAGCAACAC 5760
 DB 5701 AATTGAACGGGAGGCAACCTGCTCAACGTCGCAAGAAAGAGCGGTGCCAGCAACAC 5760
 QY 5761 ATAAGTTAAACCTTGAAGTGTTCGACACAGATTGACTAGGAGACGTAATCTACGAT 5820
 DB 5761 ATAAGTTAAACCTTGAAGTGTTCGACACAGATTGACTAGGAGACGTAATCTACGAT 5820
 QY 5821 CTTTGACACCGGTACCGCTGACGCTTGATTTTAACTGTGCAACGAGTCATGCTC 5880
 DB 5821 CTTTGACACCGGTACCGCTGACGCTTGATTTTAACTGTGCAACGAGTCATGCTC 5880
 QY 5881 GGTGTACCTGTGTGTGTCACCCGCAATCTTCAACCCCAATGGAATGTTGTTGACG 5940
 DB 5881 GGTGTACCTGTGTGTGTCACCCGCAATCTTCAACCCCAATGGAATGTTGTTGACG 5940
 QY 5941 AGTAGCGTTGATATTTTGGCGCGCTGAA 5969
 DB 5941 AGTAGCGTTGATATTTTGGCGCGCTGAA 5969

RESULT 2

AP005282/c 320550 bp DNA linear ECT 08-AUG-2002

LOCUS Corynebacterium glutamicum ATCC 13032 DNA, complete genome, section 9/10.

DEFINITION AP005282 BA000036

ACCESSION AP005282.1 GI:21325287

KEYWORDS Corynebacterium glutamicum ATCC 13032

ORGANISM Corynebacterium glutamicum ATCC 13032

SOURCE Corynebacteriaceae; Actinomycetales; Corynebacterium.

REFERENCE 1 Nakagawa, S.

AUTHORS Complete genomic sequence of Corynebacterium glutamicum ATCC 13032

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 320550)

AUTHORS Nakagawa, S.

JOURNAL Direct Submission

COMMENT Submitted (24-MAY-2002) Satoshi Nakagawa, Kyowa Hakko Kogyo Co. Ltd., Tokyo Research Laboratories; 3-6-6, Asahi-mechi, Machida, Tokyo 194-8533, Japan (E-mail: snakagawa@kxanagen.com, Tel:81-44-829-3031, Fax:81-44-813-1611)

FEATURES This sequence is conducted by collaboration of Kyowa Hakko Kogyo Co. Ltd. And Kitasato University.

location/Qualifiers

1..320550

/organism="Corynebacterium glutamicum ATCC 13032"

/mol_type="genomic DNA"

/strain="ATCC 13032"

/db_xref="taxon:196627"

/note="ATCC 13032"

188..814

/gene="G12516"

188..814

/gene="G12516"

/codon_start=1

/transl_table=1

/product="Hypothetical protein"

/protein_id="BAB99909.1"

/db_xref="GI:21325288"

[illegible]

```

/product="Ribonucleotide reductase beta subunit"
/protein_id="BAB9918.1"
/db_xref="GI:21325297"
/translation="MAADSDLSVHDAYLKEHVAIPVKAINMNSIPDSKOLEVWDELITGN
FMILPKRVSNIDIKSMGTLNEVERKAITRVFTGLTLDITQGTGASLIDPADSLHE
EAVLTNIAFMESVHAKSYNSIFMTLASTAEIINDAPRMEENIQRXKIIISYEDD
DPLKRIASVLESFLYISGFLEPMWSHSLTNITADYILILIRDEAVHGYIYIKY
OKAVAKETPERBELKEYTEFDLIDLIDYNEQYSEDLISLWTEDEVKPRFMYANKA
LNILGEGLFPADETKVPNLSLSPNADENHDFSSGSSVYIGKALENTDDMDPE"
gene
11555..12043
/gene="cg12526"
CDS
11555..12043
/gene="cg12526"
/note="PF00210:Ferritin"

Query Match      94.8%; Score 5656.2; DB 1; Length 32050;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 5793; Conservative 0; Mismatches 173; Indels 3; Gaps 3;

QY 1 AGTCGCGTGCAGCCACCATTTGATGTGTGTGTCACCGAGCTTGCGGAGGCTTTCTACATCT 60
Db 152336 AGTCGCGTGCAGCCACCATTTGATGTGTGTGTCACCGAGCTTGCGGAGGCTTTCTACATCT 152277
QY 61 AGCTTCCCTCGGCGTGGAGTGGGTTCATTAAGGGTGGATACCGCGGTGAAGTTGGG 120
Db 152276 AGCTTCCCTCGGCGTGGAGTGGGTTCATTAAGGGTGGATACCGCGGTGAAGTTGGG 152217
QY 121 GAACCATGCTGCTCTTGTGGGTTGAGGGAACGAGTGGGCTGAGAAATTTTCAAGTG 180
Db 152216 GAACCATGCTGCTCTTGTGGGTTGAGGGAACGAGTGGGCTGAGAAATTTTCAAGTG 152157
QY 181 TCTGCAGTTTTTAAATGATCATCATGCTTGGAAAGCTGAGTTAATTCAGTAACTTG 240
Db 152156 TCTGCAGTTTTTAAATGATCATCATGCTTGGAAAGCTGAGTTAATTCAGTAACTTG 152097
QY 241 CAACGACAGGCTCAAGTCCGAAGATTAATTAACCTAGATCCGTAAGCATTAAGACATCAT 300
Db 152096 CAACGACAGGCTCAAGTCCGAAGATTAATTAACCTAGATCCGTAAGCATTAAGACATCAT 152037
QY 301 CGTCTATGCTTGTGGAAGAAACCAATTAACCTCAGAAAGATGGCAAGGTGTGCTAT 360
Db 152036 CGTCTATGCTTGTGGAAGAAACCAATTAACCTCAGAAAGATGGCAAGGTGTGCTAT 151977
QY 361 ATCAAGAAAATGAGGTCAAGCATTAAGTAAATAATGAGGAAGATTTTCCCCCTCG 420
Db 151976 ATCAAGAAAATGAGGTCAAGCATTAAGTAAATAATGAGGAAGATTTTCCCCCTCG 151917
QY 421 GGGTGATGATGGCTTCTCCAACTGAAACGCGCATCATCAACGAACTCTCTGAGAAC 480
Db 151916 GGGTGATGATGGCTTCTCCAACTGAAACGCGCATCATCAACGAACTCTCTGAGAAC 151857
QY 481 CAGACCTTAATAACGAGGATTCACCCCGAACTCCCAAGATTGTTCCCGGTTTTATTG 540
Db 151856 CAGACCTTAATAACGAGGATTCACCCCGAACTCCCAAGATTGTTCCCGGTTTTATTG 151797
QY 541 ATTTTATATATCAAGGTGAAAGGTGGGCTTTCTTAAGGAAAGCGAGACCGAGCGA 600
Db 151796 ATTTTATATATCAAGGTGAAAGGTGGGCTTTCTTAAGGAAAGCGAGACCGAGCGA 151737
QY 601 GGAACACCGCGAGTATCAACCGGAACATGACGACGCGTATGTTCCAGCATGTTT 660
Db 151736 GGAACACCGCGAGTATCAACCGGAACATGACGACGCGTATGTTCCAGCATGTTT 151677
QY 661 CGGCGCGGCTGACGACATGGCAGCGCAGGTGAAAACTTATTTCTTTGTGTGAAGAG 720
Db 151676 CGGCGCGGCTGACGACATGGCAGCGCAGGTGAAAACTTATTTCTTTGTGTGAAGAG 151617
QY 721 TCCGTGCTGCGGCAATTCACCTGAGAGGCTTTCAATCAAGCATGCGTGTGGTGC 780
Db 151616 GCGTGTGCTGCGGCAATTCACCTGAGAGGCTTTCAATCAAGCATGCGTGTGGTGC 151557
QY 781 AAAACCGGATTTCAATTTTCCGCGAACCAACAGATCTTGCCCGGATGATCATGCGG 840

```

```

Db 151556 AAAACCGGATTTTCAATTTTCCGCGAACCAACAGATCTTGCCAGGTGATCATGCGG 151497
QY 841 GAAAAGTTGGATCAAAATGCATACAGATAGCGCGGAACTGACATCTTTCTGAGCTTC 900
Db 151496 GAAAAGTTGGATCAAAATGCATACAGATAGCGCGGAACTGACATCTTTCTGAGCTTC 151437
QY 901 TCGATCTCTGCGAGCGGCACACATCATTTGCTTCCCTGCGGACACATGATGACGATTTG 960
Db 151436 TCGATCTCTGCGAGCGGCACACATCATTTGCTTCCCTGCGGACACATGATGACGATTTG 151377
QY 961 ATACCATACAGCGCAATTTGCTTGGCTAAAGAAAAATGTGACCGGTCAACGCTACGC 1020
Db 151376 ATACCATACAGCGCAATTTGCTTGGCTAAAGAAAAATGTGACCGGTCAACGCTACGC 151317
QY 1021 ATTTGTTCAATGCATGCTCTCCGCTGCTGATATAGGAGATCCCGGAGCGTGGGCTTTC 1080
Db 151316 ATTTGTTCAATGCATGCTCTCCGCTGCTGATATAGGAGATCCCGGAGCGTGGGCTTTC 151257
QY 1081 TTGCTGCGGACCGGCGGCGGACATATGTTGATGATGCGGACGCGGTGATTTGG 1140
Db 151256 TTGCTGCGGACCGGCGGCGGACATATGTTGATGATGCGGACGCGGTGATTTGG 151197
QY 1141 CCGATGGAACGCTGATCTAGCTGTTCAACCAACGCTTTTTCATCAAGGACGCTATGG 1200
Db 151196 CCGATGGAACGCTGATCTAGCTGTTCAACCAACGCTTTTTCATCAAGGACGCTATGG 151137
QY 1201 AAGCGCGGAGATGCGCAAGCTGATGATCATTTTGGCGCTTTTGAACGTCACGCTCACCG 1260
Db 151136 AAGCGCGGAGATGCGCAAGCTGATGATCATTTTGGCGCTTTTGAACGTCACGCTCACCG 151077
QY 1261 ATGAGTGGCGCGTCTGCGGATGAGCGGCGCATGCGCGGCGGACACGACACTAGCGA 1320
Db 151076 ATGAGTGGCGCGTCTGCGGATGAGCGGCGCATGCGCGGCGGACACGACACTAGCGA 151017
QY 1321 GTCAATTGCTGACACAGCTGCGGAGGGATGACGCTTATGACGCGACCTTCCACACT 1380
Db 151016 GTCAATTGCTGACACAGCTGCGGAGGGATGACGCTTATGACGCGACCTTCCACACT 150957
QY 1381 CAACCGTCCCGCTAAAAATTCGCTGTCGCGATCAACGAAATGGCTAAATCCAAACCTG 1440
Db 150956 CAACCGTCCCGCTAAAAATTCGCTGTCGCGATCAACGAAATGGCTAAATCCAAACCTG 150897
QY 1441 CAAATTTTGTGATCTTGAATCAACAGCGCAGGTGCAAAAAGTCCATTTAGTATCAAG 1500
Db 150896 TAAATTTTGTGATCTTGAATCAACAGCGCAGGTGCAAAAAGTCCATTTAGTATCAAG 150837
QY 1501 TACTTTAAGTACGATAAACTATCTGAT--TTTAAAGAGTCCCAACATGGAATCACT 1559
Db 150836 TAAATTTAATACGAGCAAAAATTTCTGATTAATAAGAGATCCGACCATGAGATCATC 150777
QY 1559 ATTCGAAAGACGAGCAAGAGTGGGCAAAACAGTGGCAGTCTTAATGCAACCTTGCGC 1619
Db 150776 ATTCGAAAGACGAGCAAGAGTGGGCAAAACAGTGGCAGTCTTAATGCAACCTTGCGC 150717
QY 1619 AACTGAAAGTGAACCTTGCGGCTTGCAACAGATCTTCAACCACTGATGATCAAGAG 1679
Db 150716 AACTGAAAGTGAACCTTGCGGCTTGCAACAGATCTTCAACCACTGATGATCAAGAG 150657
QY 1679 CTCAATTCGATGATGAAGCTGGGGAAGTGTCAATTCAGAACTGCAAGGCAATTTCTGTTG 1729
Db 150656 CTCAATTCGATGATGAAGCTGGGGAAGTGTCAATTCAGAACTGCAAGGCAATTTCTGTTG 150597
QY 1729 GATGAATAGTGGGACTAACCGTGAACGATGAAAAACGCTACTTTAAACCATTCGCAAA 1799
Db 150596 GATGAATAGTGGGACTTAACCGGACGACATGAAAAACGCTACTTTAAACCATTCGTA 150537
QY 1800 GAGTTCACTGACCAATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1859
Db 150536 GAGTTCACTGACCAATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 150477
QY 1859 CCGATTCATACGAAAGACGCTGCAAGATGATGAGGCAAAAGATCCGCTGCAAGATCCGTTGAA 1919

```

Db 150476 CCTGATCCATAGAGACAGCTGAGAGTATGAGGCAAAAGATGCTGACAGATTCCTGTGAT 150417
Qy 1920 GTTCAATCTTTGGCATCGGCGGAAACGG-CACATGCTTTTCAATGAAACCATCATCTTCT 1978
Db 150416 GTTCAATCTTTGGCATCGGCGGAAACGG-CACATGCTTTTCAATGAAACCATCATCTTCT 150357
Qy 1979 CTGTAGAGCTGACAAAGGTTCAGAGGCGGTGCAACCCCTTAAACTGTGAGGACAAACGTCTGA 2038
Db 150356 CTGTAGAGCTGACAAAGGTTCAGAGGCGGTGCAACCCCTTAAACTGTGAGGACAAACGTCTGA 150297
Qy 2039 TTCTTCAACACCATCGAAGAGGTCCCAACCCACGCCGTCAACGAGGGTTTGAGCACTTGG 2098
Db 150296 TTCTTCAACACCATCGAAGAGGTCCCAACCCACGCCGTCAACGAGGGTTTGAGCACTTGG 150237
Qy 2099 TCCCGCGCGCAAAACATGCTGTTGTGSCAATCTGTGAAGAAAGCCGACCGCATCCGC 2158
Db 150236 TCCCGCGCGCAAAACATGCTGTTGTGSCAATCTGTGAAGAAAGCCGACCGCATCCGC 150177
Qy 2159 GGAACGTGTGAAGGCCAGTGAATGCTTCTTGCCCAAGTTTCAATCCTGTATGATGACAAAC 2218
Db 150176 GGAACGTGTGAAGGCCAGTGAATGCTTCTTGCCCAAGTTTCAATCCTGTATGATGACAAAC 150117
Qy 2219 ATGCCACCATCATCTGTTGATGATGAGCAGAGATATCCAACTGGAAAAAGCTGATCACTAC 2278
Db 150116 ATGCCACCATCATCTGTTGATGATGAGCAGAGATATCCAACTGGAAAAAGCTGATCACTAC 150057
Qy 2279 GCTCTCATGAGCAATTAAGCTGCGCTAAGAAACAAAGAGAAAGTACTGTGTGGGGCTA 2338
Db 150056 GCTCTCATGAGCAATTAAGCTGCGCTAAGAAACAAAGAGAAAGTACTGTGTGGGGCTA 149997
Qy 2339 TGCACACAAACATTTCCAGTTTTCGCGCGCGCATGATGATCTCTCCGACAGGAGAGC 2398
Db 149996 TGCACACAAACATTTCCAGTTTTCGCGCGCGCATGATGATCTCTCCGACAGGAGAGC 149937
Qy 2399 TCAATGATCCCAACGGAATGTAGTGAAGTACCTTCAACGTCTACTACACAGCAG 2458
Db 149936 TCAATGATCCCAACGGAATGTAGTGAAGTACCTTCAACGTCTACTACACAGCAG 149877
Qy 2459 ATCCAGGTTTCCCTTTCGACACCAAGCGCAACGCGCTGGCTCAACACCAACGCGCTTGA 2518
Db 149876 ATCCAGGTTTCCCTTTCGACACCAAGCGCAACGCGCTGGCTCAACACCAACGCGCTTGA 149817
Qy 2519 CCGGACCGGAGGATTTGAGTGAAGGACGACCTGCGCGACGCTCTTAAACCGGATGATCTCT 2578
Db 149816 CCGGACCGGAGGATTTGAGTGAAGGACGACCTGCGCGACGCTCTTAAACCGGATGATCTCT 149757
Qy 2579 ATGACCTGATGATGATCTTATCCGCTGAGAGCGTATTTACTGACGAGCACTTAACTT 2638
Db 149756 ATGACCTGATGATGATCTTATCCGCTGAGAGCGTATTTACTGACGAGCACTTAACTT 149657
Qy 2639 TCTACACCGGCAACCTTAAATTTGACGAAAGCGCGCGCAACCCAAACCTTGTGGAAG 2698
Db 149636 TCTACACCGGCAACCTTAAATTTGACGAAAGCGCGCGCAACCCAAACCTTGTGGAAG 149637
Qy 2699 TCGAGAGACCAACTGGGCTGATGAGGAGGATTCATGCGCGTTCGCTTAAATTCGCTTA 2758
Db 149636 TCGAGAGACCAACTGGGCTGATGAGGAGGATTCATGCGCGTTCGCTTAAATTCGCTTA 149577
Qy 2759 TCGACGGAACCGGACGAGGTTTCAACCCCATTTACCGGATCCATGATCAAGCCCTGATG 2818
Db 149576 TCGACGGAACCGGACGAGGTTTCAACCCCATTTACCGGATCCATGATCAAGCCCTGATG 149517
Qy 2819 GTGATGTTGGAACATGTTCTTTGGGACCACACGCGAAAACTCACCGGTGACGCGGTTT 2878
Db 149516 GTGATGTTGGAACATGTTCTTTGGGACCACACGCGAAAACTCACCGGTGACGCGGTTT 149457
Qy 2879 TATTAACCGCTGACAGATCTTTGAAAACTGGGAAATTTCTCCGTTGAAATCACTTTGAACCTA 2938
Db 149456 TATTAACCGCTGACAGATCTTTGAAAACTGGGAAATTTCTCCGTTGAAATCACTTTGAACCTA 149397
Qy 2939 GTGATGCAACACTGGTTCCTGCTCTGATCTCGTTCGATGAGCTACATGTGGAAATGCC 2998
Db 149396 GTGATGCAACACTGGTTCCTGCTCTGATCTCGTTCGATGAGCTACATGTGGAAATGCC 149337

Qy 2999 CCAACCTTTTACGCTTGCCGATGAGAAACTGGGAAAGATCTCGACGCTGATTTTCT 3058
Db 149336 CCAACCTTTTACGCTTGCCGATGAGAAACTGGGAAAGATCTCGACGCTGATTTTCT 149277
Qy 3059 GTCCAAAGATTTGAGACCGAATCCACGATGAGGTTACTCATACCAAGCTCTGACAGT 3118
Db 149276 GTCCAAAGATTTGAGACCGAATCCACGATGAGGTTACTCATACCAAGCTCTGACAGT 149217
Qy 3119 GCGGATATGTCTGACAAAGCTTGAAGAACGACCTTCGCGTTCGAGAGATTCAAGC 3178
Db 149216 GCGGATATGTCTGCGCAAGCTTGAAGAACGACCTTCGCGTTCGAGAGATTCAAGC 149157
Qy 3179 AACTGGAATTTGCGGCATGAATTTCTACGACCGCAGCTTGCAGTAAACGTTTGTATGCTT 3238
Db 149156 AACTGGAATTTGCGGCATGAATTTCTACGACCGCAGCTTGCAGTAAACGTTTGTATGCTT 149097
Qy 3239 GGCCTGAGGCTGAGATGGGGCTGCCCGGACAGATGATCAACCAACATTTGACAGAGAG 3298
Db 149096 GGCCTGAGGCTGAGATGGGGCTGCCCGGACAGATGATCAACCAACATTTGACAGAGAG 149037
Qy 3299 GATGGGTGACTGCTGACCTGACTGTGCCCGAAGCTTCAATTTGGCAACACGCGATCTAC 3358
Db 149036 GATGGGTGACTGCTGACCTGACTGTGCCCGAAGCTTCAATTTGGCAACACGCGATCTAC 148977
Qy 3359 AAGACTCTCTTCTCCCAAGGGGAGGTGGGGGTTAATCAGATCTGTATTAGTTCTGAAC 3418
Db 148976 AAGACTCTCTTCTCCCAAGGGGAGGTGGGGGTTAATCAGATCTGTATTAGTTCTGAAC 148917
Qy 3419 CTGTCCGATGACATCCGAGGCAATATTTCCCTGAGTGGAGTGGTCCGTTTGTCTG 3478
Db 148916 CTGTCCGATGACATCCGAGGCAATATTTCCCTGAGTGGAGTGGTCCGTTTGTCTG 148857
Qy 3479 TGAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3538
Db 148856 TGAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 148797
Qy 3538 ACGATATACAGCCATTGAGATATCTGACAGGATGATGACAGGTTTCACTTCGTTTCCCG 3598
Db 148796 ACGATATATACAGCCATTGAGATATCTGACAGGATGATGACAGGTTTCACTTCGTTTCCCG 148738
Qy 3599 GCCTTCAAGGATGACATATTGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 3658
Db 148737 GCCTTCAAGGATGACATATTGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 148678
Qy 3659 GTACAAATTAATTCGCAAAATCCCTTATTCGACACAAATTAACAGGTTTAAATTTGTTA 3718
Db 148677 GTACAAATTAATTCGCAAAATCCCTTATTCGACACAAATTAACAGGTTTAAATTTGTTA 148618
Qy 3719 GCTTTTGAACAAATTCATGTTCTGAATTTTCTTCCGCTTAAAGAGAAATTC 3778
Db 148617 GCTTTTGAACAAATTCATGTTCTGAATTTTCTTCCGCTTAAAGAGAAATTC 148558
Qy 3779 ATGACCAATTAAGAACTTGCGGACGATCCGCGGACATTTGGGCGGCAACATTT 3838
Db 148557 ATGACCAATTAAGAACTTGCGGACGATCCGCGGACATTTGGGCGGCAACATTT 148498
Qy 3839 GTGCGCGCGGACACATGTTGCAACGCGTTTACGCTGCTGCTCAAAAGACACCAAGATGTC 3898
Db 148497 GTGCGCGCGGACACATGTTGCAACGCGTTTACGCTGCTGCTCAAAAGACACCAAGATGTC 148438
Qy 3899 GATGCGCAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3958
Db 148437 GATGCGCAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 148378
Qy 3959 TTCCAGATCATGTCGGGCGCAGGAGATGATGATGATGATGATGATGATGATGATGATG 4018
Db 148377 TTCCAGATCATGTCGGGCGCAGGAGATGATGATGATGATGATGATGATGATGATGATG 148318
Qy 4019 ACCTCAAGACATGCTGTGTCTCAACAGACGCTCAAGATGTTGTGCTTAAACACGCC 4078
Db 148317 ACCTCAAGACATGCTGTGTCTCAACAGACGCTCAAGATGTTGTGCTTAAACACGCC 148258

QY 4079 AACTGTTCAAGCCGCTGTGTGAAGGATTTGCGGACATTTTCGTCCCGCTAATTCATC 4138
 Db 148257 AACTGTTCAAGCCGCTGTGTGAAGGATTTGCGGACATTTTCGTCCCGCTAATTCATC 148198
 QY 4139 TTGGTTGGTGGCGGCTGTCTCAATGGCTATCAACATATGTTGGTGGCGGACGATCTGTC 4198
 Db 148197 TTGGTTGGTGGCGGCTGTCTCAATGGCTATCAACATATGTTGGTGGCGGACGATCTGTC 148138
 QY 4199 GGTCCGCAATCACTGGTGAAGATGTTCCCTCAGATCAGCGGTGTGCTGAATGATCAAC 4258
 Db 148137 GGTCCGCAATCACTGGTGAAGATGTTCCCTCAGATCAGCGGTGTGCTGAATGATCAAC 148078
 QY 4259 CTGATGCAATCTGCCCGTTGCGGCTTTCTTCCAGTGTGTTGGTTCACCGCAACCAAG 4318
 Db 148077 CTGATGCAATCTGCCCGTTGCGGCTTTCTTCCAGTGTGTTGGTTCACCGCAACCAAG 148018
 QY 4319 CGTTTCGGGCGCAATGAAGTTCCTGGGCGCGGCAATGGATGAGCAATGTTCCCAAC 4378
 Db 148017 CGTTTCGGGCGCAATGAAGTTCCTGGGCGCGGCAATGGATGAGCAATGTTCCCAAC 147958
 QY 4379 CTGTTTAAACGCTAAGCACTGCGCCGCAACATGACCGCGGCGAAATGCAATGTGTCC 4438
 Db 147957 TTGGTAACGCTAAGCACTGCGCCGCAACATGACCGCGGCGAAATGCAATGTGTCC 147898
 QY 4439 CTGTTTGGTTGGATGTTGCTCAAGCTGTTTACAGGGGCAACGCTTCTGTGCTGGTG 4498
 Db 147897 CTGTTTGGTTGGATGTTGCTCAAGCTGTTTACAGGGGCAACGCTTCTGTGCTGGTG 147838
 QY 4499 GTCTCTGGATCTGGGCAACGATTCGAGAAAGTTCTGCAACGCAACGACTATGGGCACTCA 4558
 Db 147837 GTCTCTGGATCTGGGCAACGATTCGAGAAAGTTCTGCAACGCAACGACTATGGGCACTCA 147778
 QY 4559 GACTTCCTGATCAACCCAGTGTGAATCTGTGCTTCAACCGGCTCTTACGTTACTGCT 4618
 Db 147777 GACTTCCTGATCAACCCAGTGTGAATCTGTGCTTCAACCGGCTCTTACGTTACTGCT 147718
 QY 4619 ATTGATCCAGCAATGGCGGTGGGTGACTGTGCTGCAACAGGCTGAGGAGACTAT 4678
 Db 147717 ATTGATCCAGCAATGGCGGTGGGTGACTGTGCTGCAACAGGCTGAGGAGACTAT 147658
 QY 4679 GATTTCCGATGTCAGTCGCGGCTGTGCTTTCGCTGTGCTTACTCAACCAATCGTTATC 4738
 Db 147657 GATTTCCGATGTCAGTCGCGGCTGTGCTTTCGCTGTGCTTACTCAACCAATCGTTATC 147598
 QY 4739 ACTGCTGCAACGATCTTCCCGCAATGAGCTGAGCTGTTCACAGGGTGGATTC 4798
 Db 147597 ACTGCTGCAACGATCTTCCCGCAATGAGCTGAGCTGTTCACAGGGTGGATTC 147538
 QY 4799 TTCAATTCGCAACCGCATTCATGGCCATATCGCGAGGGTGAAGATGTTGGCAATG 4858
 Db 147537 TTCAATTCGCAACCGCATTCATGGCCATATCGCGAGGGTGAAGATGTTGGCAATG 147478
 QY 4859 TTCTTCCTGAGAGAGTGAAGTCAAGGGCCCTTGCAGATGCTTCAAGTGTCCGCT 4918
 Db 147477 TTCTTCCTGAGAGAGTGAAGTCAAGGGCCCTTGCAGATGCTTCAAGTGTCCGCT 147418
 QY 4919 GTTCTTGGTATTAACAGAGCTTCGATCTTCCGCTGTAACCTTCGCTGCGCTTC 4978
 Db 147417 GTTCTTGGTATTAACAGAGCTTCGATCTTCCGCTGTAACCTTCGCTGCGCTTC 147358
 QY 4979 TACATTTGATCGGTAACGAGCTATCGGTGGCGCTTGAATGCACTTTGATATCAAG 5038
 Db 147357 TACATTTGATCGGTAACGAGCTATCGGTGGCGCTTGAATGCACTTTGATATCAAG 147298
 QY 5039 GCAAGTTGCGTGGCGCTCAGAGTTTCTTGGGTGTTGTTTGAATGATCTCCAGATATG 5098
 Db 147297 GCAAGTTGCGTGGCGCTCAGAGTTTCTTGGGTGTTGTTTGAATGATCTCCAGATATG 147238
 QY 5099 GTCAATGTTCTGTTGGTGGCGGTAAGTACTTTCATGCAATGCGGCGCAAGATGCT 5158
 Db 147237 GTCAATGTTCTGTTGGTGGCGGTAAGTACTTTCATGCAATGCGGCGCAAGATGCT 147178
 QY 5159 TATGCGCTTACTTGGTGGCGCAACGGCAGCATGATGATCAACCGCTGCTCA 5218

Db 147177 TATGCGCTTACTTGGTGGCGCAACGGCAGCATGATCAAGTCAACCGCTGCTCA 147118
 QY 5219 GTGCGCTCAGGAACGACCAAGCCGGAAGCAAGACCCGAGATTTTCAAGATTC 5278
 Db 147117 GTGCGCTCAGGAACGACCAAGCCGGAAGCAAGACCCGAGATTTTCAAGATTC 147058
 QY 5279 ACCATATCCAGGCACTTTGACCGGTGAAGCTATGCACTGAGACCGTCAAGATGCC 5338
 Db 147057 ACCATATCCAGGCACTTTGACCGGTGAAGCTATGCACTGAGACCGTCAAGATGCC 146998
 QY 5339 ATGTTTGCAGCGGAAGCTTGGCTCAGGTGTTGCAATGCTGCCCAACCAAGGCGCACTG 5398
 Db 146997 ATGTTTGCAGCGGAAGCTTGGCTCAGGTGTTGCAATGCTGCCCAACCAAGGCGCACTG 146938
 QY 5399 GTTTTCAACGAGTGAACCGGAAGATGAGTGGCTTCCATCTGTGTCACGCTTTCGACATC 5458
 Db 146937 GTTTTCAACGAGTGAACCGGAAGATGAGTGGCTTCCATCTGTGTCACGCTTTCGACATC 146878
 QY 5459 CGCACTAAGGCTGAGATGTTCCATATGTGATATCTTGATGCAATTTGGTTTGCACACC 5518
 Db 146877 CGCACTAAGGCTGAGATGTTCCATATGTGATATCTTGATGCAATTTGGTTTGCACACC 146818
 QY 5519 GTAAACCTCAACGCGCAGCACTTTAACCCGCTGAAGAAAGCAGGCGATGAAGTCAAGCA 5578
 Db 146817 GTAAACCTCAACGCGCAGCACTTTAACCCGCTGAAGAAAGCAGGCGATGAAGTCAAGCA 146758
 QY 5579 GGGAGAGCTGTGTGAATTCGATATGATGATCCATTAAGCTGCAAGTTATGAGTAAAC 5638
 Db 146757 GGGAGAGCTGTGTGAATTCGATATGATGATCCATTAAGCTGCAAGTTATGAGTAAAC 146698
 QY 5639 AGCGCATGTTGTTTGAATTAACAAGAAACCGGACCTGTAACCTTACGTTTGGGC 5698
 Db 146697 AGCGCATGTTGTTTGAATTAACAAGAAACCGGACCTGTAACCTTACGTTTGGGC 146638
 QY 5699 GAAATTGAAGCGGAGCGCAACCTGCTCAACGTCGCAAGAAAGAGCGGTCCAGCAACA 5758
 Db 146637 GAAATTGAAGCGGAGCGCAACCTGCTCAACGTCGCAAGAAAGAGCGGTCCAGCAACA 146578
 QY 5759 CCATTAAGTTGAACCTTGAGTGTTCGCAACAGGTTAGCTAGGGGACGTGACTTACG 5818
 Db 146577 CCATTAAGTTGAACCTTGAGTGTTCGCAACAGGTTAGCTAGGGGACGTGACTTACG 146518
 QY 5819 ATCTTTGACACCGGATCCGATGCTTGAATTTTAAACCTGTTCAACAGATCAATGCC 5878
 Db 146517 ATCTTTGACACCGGATCCGATGCTTGAATTTTAAACCTGTTCAACAGATCAATGCC 146458
 QY 5879 TCGGTATACCTGTGTGTGTCACCCCGCAATCTTCAACCCATTTGACATGTTGTTCA 5938
 Db 146457 TCGGTATACCTGTGTGTGTCACCCCGCAATCTTCAACCCATTTGACATGTTGTTCA 146398
 QY 5939 GCAATGAGCTTGAATTTTTCGCGCGCTG 5967
 Db 146397 GCAATGAGCTTGAATTTTTCGCGCGCTG 146369

RESULT 3
 BX927155/c 349136 bp DNA linear BCT 10-JUN-2004
 LOCUS BX927155 Corynebacterium glutamicum ATCC 13032, IS fingerprint type 4-5,
 DEFINITION complete genome; segment 8/10.
 ACCESSION BX927155 BX927147
 VERSION BX927155.1 GI:41326514
 KEYWORDS complete genome.
 SOURCE Corynebacterium glutamicum ATCC 13032
 ORGANISM Corynebacterium glutamicum ATCC 13032
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE 1 [bases 1 to 349136]
 Kalinowski, J., Bathe, B., Bartels, D., Bischoff, N., Bort, M.,
 Burkowski, A., Dusch, N., Eggeling, L., Eikmanns, B. J., Gaigalat, L.,
 Goemann, A., Hartmann, M., Hutmacher, K., Kramer, R., Linke, B.,
 McHardy, A. C., Meyer, F., Mockel, B., Pfeifferle, W., Puhler, A.,


```

gene      complement(8387. .8911)
          /gene="dctx"
          /locus_tag="cg2569"
          /complement(8387. .8911)
CDS
          /gene="dctx"
          /locus_tag="cg2569"
          /codon_start=1
          /transl_table=11
          /product="PUTATIVE C4-DICARBOXYLATE TRANSPORT SYSTEM
          (PERMEASE SMALL PROTEIN)"
          /protein_id="CAF21004.1"
          /db_xref="GI:41326522"
          /translation="MMNFKSIYCVTAMOVFSROVLSHSPSTWSEELFYWLSPAGS
          APLFERGHIADVPIARKLPVSAORLVQIVQLLVPAIIGIMIGVLAASIANQO
          LTALEPLTGMVYVVIPIAGVFIALFAIIDLIEVAKKEPYPVDESEPPDDLEEA
          OSADISASSABERN"
gene      complement(8908. .9657)
          /gene="dctp"
          /locus_tag="cg2570"
          /complement(8908. .9657)
CDS
          /gene="dctp"
          /locus_tag="cg2570"
          /codon_start=1
          /transl_table=11
          /product="PUTATIVE C4-DICARBOXYLATE-BINDING PROTEIN"
          /protein_id="CAF21005.1"
          /db_xref="GI:41326523"
          /translation="MMGSEVERKRYLRSPDFMPTPEVGVIGSIHEOMSVIRDDMEPL
          FOSPLAESDNISVIGFPTGTRNLVTTDAVVRKPADLAKKLRVQSAWHIRIETLMSG
          SATPLTGEVITYAMGSGVLDGAENNEISVYONHFEVARVNSNTNHLGLDYMVRHD
          LLDMASEPDRLEFLEEMDAMATEHTDMLNTEFDVAIERAKGSAFVEVDAQFTDAL
          APIKDEFLTSEFRELVEAVRAADTSGGAAAS"
gene      complement(9722. .11569)
          /gene="1epa"
          /locus_tag="cg2571"
          /complement(9722. .11569)
CDS
          /gene="1epa"
          /locus_tag="cg2571"
          /complement(9722. .11569)
          /locus_tag="cg2571"
          /codon_start=1

Query Match      94.8%; Score 5656.2; DB 1; Length 349136;
Beet Local Similarity 97.1%; Pred. No. 0;
Matches 5793; Conservative 0; Mismatches 173; Indels 3; Gaps 3;

```

```

Db      343974 GGGTATGATGGCTTTCTCCAACTCGAAGAACGGGATCATACGGGAACCTCTCGAGAAC 343915
Qy      441 CAGCAGCTAAAAACGAGGATTCACCCCGAACTCCCAAGATTTGTTCCCGTTTATTTG 540
Db      343914 CAGCAGCTAAAAACGAGGATTCACCCCGAACTCCCAAGATTTGTTCCCGTTTATTTG 343855
Qy      541 ATCTTCATATACGGGTGGAACCGGTGGCGGTTTCTACGGGAACGAGGACGAGCGGA 600
Db      343854 ATCTTCATATACGGGTGGAACCGGTGGCGGTTTCTACGGGAACGAGGACGAGCGGA 343795
Qy      601 GGAACACCGCGCAGTATACCGCGAATGGCAACGACCGGTGATGGCAAGCATGTTT 660
Db      343794 GGAATGCCCGCAGTATACCGCGAATGGCAACGACCGGTGATGGCAAGCATGTTT 343735
Qy      661 CGGCGCCGCTGACCGCATGGCGAGCGAGGTGGAACCTTATTCCTTGTGTGAAGAG 720
Db      343734 CGGCGCCGCTGACCGCATGGCGAGCGAGGTGGAACCTTATTCCTTGTGTGAAGAG 343675
Qy      721 TCTGCTGTGGCGCATTCACCTCGAGGGCCCTTTCATCAACGATGCCGTTGTGGTCTC 780
Db      343674 GCTGCTGTGGCGCATTCACCTCGAGGGCTCTTTCATCAACGATGCCGTTGTGGTCTC 343615
Qy      781 AAAACCCGGAATTCATTTTCCCGGCAACCGAAGATCTTGCCCGGATGATCCATGCGG 840
Db      343614 AAAACCCGGAATTCATTTTCCCGGCAACCGAAGATCTTGCCCGGATGATCCATGCGG 343555
Qy      841 GAAAAGTTGGATCAATTCATACAGTACAGGCGCGGAACCTGACAACTTTCTGAGCTTC 900
Db      343554 GAAAAGTTGGATCAATTCATACAGTACAGGCGCGGAACCTGACAACTTTCTGAGCTTC 343495
Qy      901 TCGATCTGTGGCGAGCGCACCATCATTCCTTCTGCGGCAACATGATCAGATTTTG 960
Db      343494 TCGATCTGTGGCGAGCGCACCATCATTCCTTCTGCGGCAACATGATCAGATTTTG 343435
Qy      961 ATACCACTACAGCGCAATTCCTTGGCTAAAGAAAATGTGACGATCAACGCTACGC 1020
Db      343434 ATACCACTACAGCGCAATTCCTTGGCTAAAGAAAATGTGACGATCAACGCTACGC 343375
Qy      1021 ATTTGTTCAATGCCATGCTCTCCGTCATCATATAGGCTCCCGGCGAGCGTGGCGCTTTC 1080
Db      343374 ATTTGTTCAATGCCATGCTCTCCGTCATCATATAGGCTCCCGGCGAGCGTGGCGCTTTC 343315
Qy      1081 TTGCTGCGGCAAGTCCCGGAGACGATATGTTGATGTCGCGACGCGGTGCAATTTTG 1140
Db      343314 TTGCTGCGGCAAGTCCCGGAGACGATATGTTGATGTCGCGACGCGGTGCAATTTTG 343255
Qy      1141 CCGATGAAAGGTGATCTGCTGTTCCAAACACGCTTTTTCATCAACGACGCTATGG 1200
Db      343254 CCGATGAAAGGTGATCTGCTGTTCCAAACACGCTTTTTCATCAACGACGCTATGG 343195
Qy      1201 AAGCGCGGGAATGCAACGATGATCATTTTGGGGGTTTGAAGTCAACCGTACCG 1260
Db      343194 AAGCGCGGGAATGCAACGATGATCATTTTGGGGGTTTGAAGTCAACCGTACCG 343135
Qy      1261 ATGAGTCCGCGTCTGCGCATGGCGGCGCATCGCGGGGCAACGACACTAGGGA 1320
Db      343134 ATGAGTCCGCGTCTGCGCATGGCGGCGCATCGCGGGGCAACGACACTAGGGA 343075
Qy      1321 GTCAAGTGTGACACGCTGCGGAGGGGTATGACGCTTATGACGCGACCTCCACACT 1380
Db      343074 GTCAAGTGTGACACGCTGCGGAGGGGTATGACGCTTATGACGCGACCTCCACACT 343015
Qy      1381 CAACCGTCCGCGCTAAATTCCTGCGTCTGCGCATGCAAGAAATGCTAAATCCAAACCTG 1440
Db      343014 CAACCGTCCGCGCTAAATTCCTGCGTCTGCGCATGCAAGAAATGCTAAATCCAAACCTG 342955
Qy      1441 CAATTTTGTGATCTTACCAACGCGCAGGTGCAAAAGTCCATTTAGTATCAAG 1500
Db      342954 TAAATTTTGTGATCTTACCAACGCGCAGGTGCAAAAGTCCATTTAGTATCAAG 342895
Qy      1501 TACTTTAAGTACGATAAACTATCTTGAT-TTTAAAGAGTCCCAACATGAAATCACT 1559

```


Db 342894 TAATTAAATACGAGCAAACTTTCGTGATTAATAAAGAGTCCGACATGACATCATC 342835
QY 1560 ATCTGAAAGACGACGAAAGTCGGCAAGCGTTGCACTCTTAATCGCACTCTTGGCC 1619
Db 342834 ATCTGAAAGACGACGAAAGTCGGCAAGCGCGCGACCTCTGATCGCACTCTTGGCA 342775
QY 1620 AACAAAGGTGAACTTTGGGGCTTGGCAAGATCTCTACCACTGAGTACTCAAGAG 1679
Db 342774 ACTAAGGGGGAACCTTTGGGGCTTGGCACTGATCTCTACCTTTGAGCACTCAAGAG 342715
QY 1680 CTCATTCGATGTATGAAAGCTGGGGAAGTCTCATTTCAAGAACTGCAAGGCAATCTTGTG 1739
Db 342714 CTCATTCGATGTATGAAAGCTGGGGAAGTCTCATTTCAAGAACTGCAAGGCAATCTTGTG 342655
QY 1740 GATGAAATACGTGGGACTTAACCCGTGACGATGAAAAAGAGTACTTTAAACATTCGCCAA 1799
Db 342654 GATGAAATACGTGGGACTTAACCCGTGACGATGAAAAAGAGTACTTTAAACATTCGCCAA 342595
QY 1800 GAGTTCATGACCACTGACATCGTTGATGAAAGAGTCTTACAGCCGAGATGGTGGCAAC 1859
Db 342594 GAGTTCATGACCACTGACATCGTTGATGAAAGAGTCTTACAGCCGAGATGGTGGCAAC 342535
QY 1860 CCTGATCCATACGAGACAGCTGAGATGAGGCAAAAGATCGCTGCAAGATCCGTTGAA 1919
Db 342534 CCTGATCCATACGAGACAGCTGAGATGAGGCAAAAGATCGCTGCAAGATCCGTTGAT 342475
QY 1920 GTTCAATCTCTTGGCATCGGCGGAAACG- CACATGCTTTCACTTGAACATATCTTCT 1978
Db 342474 GTTCAATCTCTTGGCATCGGCGGAAACGCGCACATGCTTTCACTTGAACATATCTTCT 342415
QY 1979 CTGTGAGACTGACAAAGGTCGAGGCGTGCACCTTAAACGTGAGGCAACGCTGCA 2038
Db 342414 CTGTGAGACTGACAAAGGTCGAGGCGTGCACCTTAAACGTGAGGCAACGCTGCA 342355
QY 2039 TTCTTCAACACCATGCAAGAGTCCCAACCAACGCGTCAACCCAGGGTTTGGCACTTTG 2098
Db 342354 TTCTTCAACACCATGCAAGAGTCCCAACCAACGCGCTCAACCCAGGGTTTGGCACTTTG 342295
QY 2099 TCCCGCGCCCAAAACATCGTGTGTGGGCACTGTGAAAGAAAGCCGACGCAATCCG 2158
Db 342294 TCCCGCGCCCAAAACATCGTGTGTGGGCACTGTGAAAGAAAGCCGACGCAATCCG 342235
QY 2159 GGAACGTGGAAGGCCAGTGACTGCTTTCGCGAGTTCCATCTCTGATGATGACAAAC 2218
Db 342234 GGAACGTGGAAGGCCAGTGACTGCTTTCGCGAGTTCCATCTCTGATGATGACAAAC 342175
QY 2219 ATGCCACATCATCTGTGATGAGGACGAGATATCCAGCTGGAACCGCTGATCACTAC 2278
Db 342174 ATGCCACATCATCTGTGATGAGGACGAGATATCCAGCTGGAACCGCTGATCACTAC 342115
QY 2279 CGTTCATGAGCAATTAAGCTGCGCTTGAAGAAAGAAAGTACTGTGTGGGGCTA 2338
Db 342114 CGTTCATGAGCAATTAAGCTGCGCTTGAAGAAAGAAAGTACTGTGTGGGGCTA 342055
QY 2339 TGCACACAGAACTTTCCAGTTTGCCTGCGCTGCGATGATGATCTCTCCGACGGGCGAG 2398
Db 342054 TGCACACAGAACTTTCCAGTTTGCCTGCGCTGCGATGATGATCTCTCCGACGGGCGAG 341995
QY 2399 TCAATGATCCAAACGGAATGTACGTGATGAGATACCTTCCACGTCTACTACACAGAG 2458
Db 341994 TCAATGATCCAAACGGAATGTACGTGATGAGATACCTTCCACGTCTACTACACAGAG 341935
QY 2459 ATCCAGGTTTCCCTTGCACCAAGCGCAACGCGCTGGGCTCACACACACGCGCTTGA 2518
Db 341934 ATCCAGGTTTCCCTTGCACCAAGCGCAACGCGCTGGGCTCACACACACGCGCTTGA 341875
QY 2519 CCGGACCGGACGATTTGAGTGGAGCAGCTGCGCGACGCTCTTAAACCGGATGATCTCT 2578
Db 341874 CCGGACCGGACGATTTGAGTGGAGCAGCTGCGCGACGCTCTTAAACCGGATGATCTCT 341815
QY 2579 ATGACCTGATGATGATCTATTCGCGTGAAGCGGATTTACTGACGACACTTAAACTTT 2638
Db 341814 ATGACCTGATGATGATCTATTCGCGTGAAGCGGATTTACTGACGACACTTAAACTTT 341755

QY 2639 TCTACACCGGCAACCTTAAATTTGACGGAAGCGCGCCGACCCCAAACTGTGTGAAG 2698
Db 341754 TCTACACCGGCAACCTTAAATTTGACGGAAGCGCGCCGACCCCAAACTGTGTGAAG 341695
QY 2699 TCGAGACCCCACTGGGCTGATGAGGCGGATTCATGCGCTTGCCTTAAATTCGCTTAA 2758
Db 341694 TCGAGACCCCACTGGGCTGATGAGGCGGATTCATGCGCTTGCCTTAAATTCGCTTAA 341635
QY 2759 TCGAGGACCCGCGACGCGTTTCAACCCCATTTACCGGATCCATGATCAGCCTGATG 2818
Db 341634 TCGAGGACCCGCGACGCGTTTCAACCCCATTTACCGGATCCATGATCAGCCTGATG 341575
QY 2819 GTGATGTTGGAACATGTTCTTTGGGGCCCAACGCGAAACCTCACCGGTGACGCGTTC 2878
Db 341574 GTGATGTTGGAACATGTTCTTTGGGGCCCAACGCGAAACCTCACCGGTGACGCGTTC 341515
QY 2879 TATACCGCTGACAGATCTTGAAGAACTGGGAATTCGCGGTGAATCACTTTGACCTCA 2938
Db 341514 TATACCGCTGACAGATCTTGAAGAACTGGGAATTCGCGGTGAATCACTTTGACCTCA 341455
QY 2939 GTGATGACAACTGGTTCTGCTCTGATCTGCTTCCGATGAGCTAATGTGGGAATGCC 2998
Db 341454 GTGATGACAACTGGTTCTGCTCTGATCTGCTTCCGATGAGCTAATGTGGGAATGCC 341395
QY 2999 CCAACTTTTACGCTTCCGCGATGAAAGAACTGGGCAAGATCTCCGCTGCTGATTTTCT 3058
Db 341394 CCAACTTTTACGCTTCCGCGATGAAAGAACTGGGCAAGATCTCCGCTGCTGATTTTCT 341335
QY 3059 GTCCAAAGATTTGACCCGATTCACGATGAGGTTACTACTACGCACTCTGACCACT 3118
Db 341334 GTCCAAAGATTTGACCCGATTCACGATGAGGTTACTACTACGCACTCTGACCACT 341275
QY 3119 GCGGATATGTCGTGCAAAAGCTTGAAGAAAGCACTTCCGCTTTGAGAGATTCAGCG 3178
Db 341274 GCGGATATGTCGTGCAAAAGCTTGAAGAAAGCACTTCCGCTTTGAGAGATTCAGCG 341215
QY 3179 AGCTGATTTGGGCGCATGAAATCTTACGACACCGCAAGTTGCAATTAACGTTTCAATGCT 3238
Db 341214 AGCTGATTTGGGCGCATGAAATCTTACGACACCGCAAGTTGCAATTAACGTTTCAATGCT 341155
QY 3239 GCGTGGGGCTGATGAGGGGCTGCCGCGCAGAGATGATCAACCAAGTTGCAAGAG 3298
Db 341154 GCGTGGGGCTGATGAGGGGCTGCCGCGCAGAGATGATCAACCAAGTTGCAAGAG 341095
QY 3299 GATGGGTGACCTGCTGACTGTGCTCCGCAACCTTCAATTTGCGCAACACGCGATCTAC 3358
Db 341094 GATGGGTGACCTGCTGACTGTGCTCCGCAACCTTCAATTTGCGCAACACGCGATCTAC 341035
QY 3359 AAGAGCTCTTCTCCCAAGGGGAGTGGGGGATATCAAGTCTGATTAAGTTGTAAC 3418
Db 341034 AAGAGCTCTTCTCCCAAGGGGAGTGGGGGATATCAAGTCTGATTAAGTTGTAAC 340975
QY 3419 CTGTCCGATGACATTCGAGGCAATATTTCCCTCGAGTGGAGTGGTCCGTTTGTCTG 3478
Db 340974 CTGTCCGATGACATTCGAGGCAATATTTCCCTCGAGTGGAGTGGTCCGTTTGTCTG 340915
QY 3479 TGAATCGTATGATGATCTGCGCTGAGTGAAGTGAACCTGCGCAATTAAGTATGATCGG 3538
Db 340914 TGAATCGTATGATGATCTGCGCTGAGTGAAGTGAACCTGCGCAATTAAGTATGATCGG 340855
QY 3539 ACGATTAATACCACTTGAATTAATCTGCAAGTGAATGAGACAGTTTCACTTTCGCG 3598
Db 340854 ACGATTAATACCACTTGAATTAATCTGCAAGTGAATGAGACAGTTTCACTTTCGCG 340796
QY 3599 GCGTCAAAAGTGAACATTAATGAGATTAAGTGAATTAAGAGGCTTTTGTGGCAAT 3658
Db 340795 GCGTCAAAAGTGAACATTAATGAGATTAAGTGAATTAAGAGGCTTTTGTGGCAAT 340736
QY 3659 GTACAAATATCTTGGCAAAATCCCTTGAATGCAACAAATTAACAGGTTTAATATGTTTA 3718
Db 340735 GTACAAATATCTTGGCAAAATCCCTTGAATGCAACAAATTAACAGGTTTAATATGTTTA 340676

QY 3719 GCTTTGAAACAATTCATGCTGAATATTTTGTCTTCCTCCGCTTAAGGAAATTC 3778
 Db 340675 GCTTTGAAACAATTCATGCTGAATATTTTGTCTTCCTCCGCTTAAGGAAATTC 340616
 QY 3779 ATGGACCAATTAAGACCTCGCGGACAGCATCCGCGGACATTGGGGGGAGAGACAATTC 3838
 Db 340615 ATGGACCAATTAAGACCTCGCGGACAGCATCCGCGGACATTGGGGGGAGAGACAATTC 340556
 QY 3839 GTCCGCGCGCACACCTGTGCAACGCGTTTACGCTCTGTGCTCAAGACAACCAAGATGTG 3898
 Db 340555 GTCCGCGCGCACACCTGTGCAACGCGTTTACGCTCTGTGCTCAAGACAACCAAGATGTG 340496
 QY 3899 GATCCGCAAAAGTCTGATATGATCAATCTGAAAGGACCTTTGAAACTGGCGGATG 3958
 Db 340495 GATCCGCAAAAGTCTGATATGATCAATCTGAAAGGACCTTTGAAACTGGCGGATG 340436
 QY 3959 TTCCAGATCATGTCGGGCGCAGGCGATGTGATCAATGTTTCAAGAACTCATGACGA 4018
 Db 340435 TTCCAGATCATGTCGGGCGCAGGCGATGTGATCAATGTTTCAAGAACTCATGACGA 340376
 QY 4019 ACCTCGCAAGACATGCTGTGTCTGCAAGACAGCTCAAAAGATGTTGTGCTAACAACGCC 4078
 Db 340375 ACCTCGCAAGACATGCTGTGTCTGCAAGACAGCTCAAAAGATGTTGTGCTAACAACGCC 340316
 QY 4079 AACTGCTTCAAGCCGTGCTGTGAAGGTAATGGGGGACATTTTGTCCCGCTGATCCATTC 4138
 Db 340315 AACTGCTTCAAGCCGTGCTGTGAAGGTAATGGGGGACATTTTGTCCCGCTGATCCATTC 340256
 QY 4139 TTGGTTGTGCGGCTGTGCTCAATGCTATCAACAATGTTGTTGCGAGGATCTGTTC 4198
 Db 340255 TTGGTTGTGCGGCTGTGCTCAATGCTATCAACAATGTTGTTGCGAGGATCTGTTC 340196
 QY 4199 GGTCCGCAATCACTGCTGTGAAGTGTCTCCAGATCAAGCGGTGTGCTGAATGATCAAC 340136
 Db 340195 GGTCCGCAATCACTGCTGTGAAGTGTCTCCAGATCAAGCGGTGTGCTGAATGATCAAC 340136
 QY 4259 CTGATGGCAATGCGCGCGTTGCGCTCTTGCCAGCTGTTGTTGCTTCCGCAACCAAG 4318
 Db 340135 CTGATGGCAATGCGCGCGTTGCGCTCTTGCCAGCTGTTGTTGCTTCCGCAACCAAG 340076
 QY 4319 CGTTTCGTGCGCAATGATGTTCTCTGCGCGCGCGCATTTGATGCGATGTTTCCCAAC 4378
 Db 340075 CGTTTCGTGCGCAATGATGTTCTCTGCGCGCGCGCATTTGATGCGATGTTTCCCAAC 340016
 QY 4379 CTGCTTAAAGGCTACGACGTCGCGCGCACATGACCGCGCGGGAATGCAATGTTGTC 4438
 Db 340015 CTGCTTAAAGGCTACGACGTCGCGCGCACATGACCGCGCGGGAATGCAATGTTGTC 339956
 QY 4439 CTGTTTGTGTTGATGTTGCTCAAGCTGTTACAGGGACCGTGTCTCTGTGCTGTG 4498
 Db 339955 CTGTTTGTGTTGATGTTGCTCAAGCTGTTACAGGGACCGTGTCTCTGTGCTGTG 339896
 QY 4499 GTCTCTGGAATTCGCAACGATCGAAGATTCCTGCAACAGGACTCATGCGGACCTGCA 4558
 Db 339895 GTCTCTGGAATTCGCAACGATCGAAGATTCCTGCAACAGGACTCATGCGGACCTGCA 339836
 QY 4559 GACTTCTGATCAACCCAGTGTGACTCTGCTGCTCAACCGGCTCTCTTAAGTTCAATGCT 4618
 Db 339835 GACTTCTGATCAACCCAGTGTGACTCTGCTGCTCAACCGGCTCTCTTAAGTTCAATGCT 339776
 QY 4619 ATTGCTCAAGCAATGCGTGGTGGTGAATGCTGTGCAACAGCGTCTGACGGGACTCTAT 4678
 Db 339775 ATTGCTCAAGCAATGCGTGGTGGTGAATGCTGTGCAACAGCGTCTGACGGGACTCTAT 339716
 QY 4679 GATTTGCGTGGTTCAGTGGCGGTCTGCTTTTGTGCTGTGCTTACTCAACATGTTATC 4738
 Db 339715 GATTTGCGTGGTTCAGTGGCGGTCTGCTTTTGTGCTGTGCTTACTCAACATGTTATC 339656
 QY 4739 ACTGCTTGAACAGTCTTCCCGCAATGAGCTGGAGCTGTCAACAGGGTGAATCC 4798
 Db 339655 ACTGCTTGAACAGTCTTCCCGCAATGAGCTGGAGCTGTGTCAACAGGGTGAATCC 339596
 QY 4799 TTCACTTTCGCAACCGCATCATGCGCAATATGCGCGAGGGTCAAGCATGTTTGGCAGTG 4858

Db 339595 TTCACTTTCGCAACCGCATCATGCGCAATATGCGCGAGGGTGGGATGTTTGGCAGTG 339536
 QY 4859 TTCTTCTTACGCAAGACTGAAAGCTCAAGGCGCTTCAGGTGCTTCAAGTGTCTCCGCT 4918
 Db 339535 TTCTTCTTACGCAAGACTGAAAGCTCAAGGCGCTTCAGGTGCTTCAAGTGTCTCCGCT 339476
 QY 4919 GTTCTTGTATTAAGAGCTTGGGATGCTTGGTGTGAACCTTGGCTGGCGTGGCGTTC 4978
 Db 339475 GTTCTTGTATTAAGAGCTTGGGATGCTTGGTGTGAACCTTGGCTGGCGTGGCGTTC 339416
 QY 4979 TACATTTGGATCGGTACCGCAGCTATGCTGCGCTTGTGATGCACTTTGATATCAAG 5038
 Db 339415 TACATTTGGATCGGTATCGGACGATGCTGCTGCGCTTGTGATGCACTTTATATCAAG 339356
 QY 5039 GCAGTGTGCTGGCGCTGCGAGTTTCTTGGGTGTGTTTCTATGATGCTCCGATATG 5098
 Db 339355 GCAGTGTGCTGGCGCTGCGAGTTTCTTGGGTGTGTTTCTATGATGCTCCGATATG 339296
 QY 5099 GTCAATGTTCTGTTTGGCGGATGTTACCTTGTCAATGCTGCGGCGAGCATGCT 5158
 Db 339295 GTCAATGTTCTGTTTGGGTGTGCAATGTTTACCTTCTTCAATGCTGCGGCGAGCATGCT 339236
 QY 5159 TATGCGCTTTTACTTGTGTTGCGCGCAACGCGACATGATCCAGATGCAACCGCTGCTCA 5218
 Db 339235 TATGCGCTTTTACTTGTGTTGCGCGCAACGCGACATGATCCAGATGCAACCGCTGCTCA 339176
 QY 5219 GTGCTGTGAGGAGAGACCAAGACCGGAGCAGAAACCCCGGAAATTTTCAAGATTC 5278
 Db 339175 GTGCTGTGAGGAGAGACCAAGACCGGAGCAGAAACCCCGGAAATTTTCAAGATTC 339116
 QY 5279 ACCATCATCGAGGACCTTTGACCGGTGAACCTATGCACTGACGACGTCAGCGATGCC 5338
 Db 339115 ACCATCATCGAGGACCTTTGACCGGTGAACCTATGCACTGACGACGTCAGCGATGCC 339056
 QY 5339 ATGTTTGCACGCGAAACCTTGGCTCAGTGTGCGATGTTCCCACTCAAGGGGACGTC 5398
 Db 339055 ATGTTTGCACGCGAAACCTTGGCTCAGTGTGCGATGTTCCCACTCAAGGGGACGTC 338996
 QY 5399 GTTTCACAGTGAAGCGGAAAGATGCTGTGCGCTTCCCATGCTGATCGCTTGGCAGTC 5458
 Db 338995 GTTTCACAGTGAAGCGGAAAGATGCTGTGCGCTTCCCATGCTGATCGCTTGGCAGTC 338936
 QY 5459 GCACTTAAAGGCTGAGAGATGTTCCATATGATGATATCTTGATGCAATTTGTTGACACC 5518
 Db 338935 GCACTTAAAGGCTGAGAGATGTTCCATATGATGATATCTTGATGCAATTTGTTGACACA 338876
 QY 5519 GTTAACTTCAACGCGACGCACTTTAACCCGCTGAAAGACAGGGGATGAACTCAAGCA 5578
 Db 338875 GTTAACTTCAACGCGACGCACTTTAACCCGCTGAAAGACAGGGGATGAACTCAAGCA 338816
 QY 5579 GGGGAGCTGCTGTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5638
 Db 338815 GGGGAGCTGCTGTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 338756
 QY 5639 ACGCGATTTGTTTGTGCAATTAACAAGAAACCGGACCTGTAAACACTTACGTTGGGC 5698
 Db 338755 ACGCGATTTGTTTGTGCAATTAACAAGAAACCGGACCTGTAAACACTTACGTTGGGC 338696
 QY 5699 GAAATTTGAAGCGGAGCCAACTGCTCAACGTCGCAAGAAAGAAAGGCGGTGCAACA 5758
 Db 338695 GAAATTTGAAGCGGAGCCAACTGCTCAACGTCGCAAGAAAGAAAGGCGGTGCAACA 338636
 QY 5759 CCATAAGTTGAACCTTGAAGTTGCAACAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTT 5818
 Db 338635 CCATAAGTTGAACCTTGAAGTTGCAACAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTT 338576
 QY 5819 ATCTTTGACACCGGATACCGTACGCTTGAAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTT 5878
 Db 338575 ATCTTTGACACCGGATACCGTACGCTTGAAGTTTGAAGTTTGAAGTTTGAAGTTTGAAGTT 338516
 QY 5879 TCGGTGTACCTGTGTGTGTCACCCCGCAATTTTCAACCCCATTTGGACATGTTGTTCA 5938

RESULT 4
AX127152/c 349980 bp DNA linear PAT 11-MAY-2001
LOCUS AX127152
DEFINITION Sequence 7068 from Patent EP1108790.
ACCESSION AX127152
VERSION AX127152.1 GI:14041140
KEYWORDS
SOURCE
ORGANISM
Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteriales;
Corynebacteriaceae; Corynebacterium.
REFERENCE
AUTHORS
1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayashi, M., Ochiai, K.,
Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.,
TITLE Novel polynucleotides
JOURNAL Patent: EP 1108790-A 7068 20-JUN-2001;
KYOWA HAKKO KOGYO CO., LTD. (JPN)
FEATURES
source
1. 349980
/organism="Corynebacterium glutamicum"
/mol_type="genomic DNA"
/db_xref="taxon:11718"
/note="Seq 1 to long (3,309,400) split in 11, seq 7068
2,700,001 3,049,980"
ORIGIN
Query Match 94.8%; Score 5656.2; DB 6; Length 349980;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 5793; Conservative 0; Mismatches 173; Indels 3; Gaps 3;
QY 1 AGTCGCTGACGCGACCATTTGATGTGTGTCACGAGCTTGGGAGGCTTTCTATCAT 60
DB 115585 AGTCGCTGACGCGACCATTTGATGTGTGTCACGAGCTTGGGAGGCTTTCTATCAT 115526
QY 61 AGCTCCCGCTCGGCGTGTGAGTGGGTCAATTACGGGTGGATCACCGCGGTGAAGTTGGC 120
DB 115525 ACACGTGCGGTGCGGCGTGTGAGTGGGTCAATTACGGGTGGATCACCGCGGTGAAGTTGGC 115466
QY 121 GAACCATGTGTCTCTGTGTGTGTGAGGAAAGAGTGGGTGAGAGTTTTCAGAGTG 180
DB 115465 GAACCATGTGTCTCTGTGTGTGTGAGGAAAGAGTGGGTGAGAGTTTTCAGAGTG 115406
QY 181 TCTGAGATTTTAAGTTATGATCATCATCTGGAAGGCTGAGATTAATCAGATGACCTG 240
DB 115405 TCTGAGATTTTAAGTTATGATCATCATCTGGAAGGCTGAGATTAATCAGATGACCTG 115346
QY 241 CAACAGCAGGCTCAAGTCCGAAGATTAATTAATCTAGATCCGTAGACATAGACATCAT 300
DB 115345 CAACAGCAGGCTCAAGTCCGAAGATTAATTAATCTAGATCCGTAGACATAGACATCAT 115286
QY 301 CGTCTATCTGTGTGTGGAAGAAACCAATTAATCTGAAAGATGCGAAGTGTGTCATT 360
DB 115285 CGTCTATCTGTGTGTGGAAGAAACCAATTAATCTGAAAGATGCGAAGTGTGTCATT 115226
QY 361 ATCAAGAAATGAGGTCAAGCATTAAGAAATTAAGGAGAAATGTTTCCCTCCCTG 420
DB 115225 ATCAAGAAATGAGGTCAAGCATTAAGAAATTAAGGAGAAATGTTTCCCTCCCTG 115166
QY 421 GGGTATGATGCTTTCTCAACTCGAAACGGCATCATCAAGAACTCTCTGAGAAC 480
DB 115165 GGGTATGATGCTTTCTCAACTCGAAACGGCATCATCAAGAACTCTCTGAGAAC 115106
QY 481 CAGCATTAAGAAACGAGATTCACCCGAACTCCCGCAGATGTTTCCCGGTTTATTG 540
DB 115105 CAGCATTAAGAAACGAGATTCACCCGAACTCCCGCAGATGTTTCCCGGTTTATTG 115046

QY 541 ATCTTCAATATCAGGTGGAACGGTGGCGGCTTCTTCAAGGGAACGAGAGCAAGGCGA 600
DB 115045 ATCTTCAATATCAGGTGGAACGGTGGCGGCTTCTTCAAGGGAACGAGAGCAAGGCGA 114986
QY 601 GGAACACCGCGAGATATCACCGCGAACAATGACGACGCGTGTATGTTCCAGACATGTTT 660
DB 114985 GGAATGCGCGAGATATCACCGCGAACAATGACGACGCGTGTATGTTCCAGACATGTTT 114926
QY 661 CGGCGCGGCTGACGACATGCGAGCGGAGTGGAAACCTTATTCCTTGTGTGAAGAG 720
DB 114925 CGGCGCGGCTGACGACATGCGAGCGGAGTGGAAACCTTATTCCTTGTGTGAAGAG 114866
QY 721 TCTGTGTGCGGAGATTAATCACTGAGGCGCTTATCAAGGACATGCGGTGGGCTTC 780
DB 114865 GCTGTGTGCGGAGATTAATCACTGAGGCGCTTATCAAGGACATGCGGTGGGCTTC 114806
QY 781 AAAACCGGATTTTATTTTCCCGGCAACCAACAGATCTTCCGCGGTATCATATGCGG 840
DB 114805 AAAACCGGATTTTATTTTCCCGGCAACCAACAGATCTTCCGCGGTATCATATGCGG 114746
QY 841 GAAAAGTTGATCAATGATCATAGTACGAGCGCGGAACTGACATCTTTGTAGCTTC 900
DB 114745 GAAAAGTTGATCAATGATCATAGTACGAGCGCGGAACTGACATCTTTGTAGCTTC 114686
QY 901 TCGATCTGCGGAGGACCAATCATTTGCTTCTTGGGACACTGATGAGATTTTG 960
DB 114685 TCGATCTGCGGAGGACCAATCATTTGCTTCTTGGGACACTGATGAGATTTTG 114626
QY 961 ATACCACTACACGCGCAATTTGCTTGGCTAAAGAAAATGTCAGCGTACGCGTACGC 1020
DB 114625 ATACCACTACACGCGCAATTTGCTTGGCTAAAGAAAATGTCAGCGTACGCGTACGC 114566
QY 1021 ATTTGTTCAATGATGCTCTCGGTCATCATAGGCTCCCGGACGCTGGCGCTTTC 1080
DB 114565 ATTTGTTCAATGATGCTCTCGGTCATCATAGGCTCCCGGACGCTGGCGCTTTC 114506
QY 1081 TTGCTGCGGACATGCGCGGAGCGATATGTTGATGATGCGAGCGCGTGAATTGG 1140
DB 114505 TTGCTGCGGACATGCGCGGAGCGATATGTTGATGATGCGAGCGCGTGAATTGG 114446
QY 1141 CCGATGGAACGCTGATAGCTCGTTCCAAACGCTTTTTCATCAAGGACGACATG 1200
DB 114445 CCGATGGAACGCTGATAGCTCGTTCCAAACGCTTTTTCATCAAGGACGACATG 114386
QY 1201 AAGCGCGGAAATGCGAAGCGGTGATCATTTTGGGCTTTGAACGTCAACCTGACCG 1260
DB 114385 AAGCGCGGAAATGCGAAGCGGTGATCATTTTGGGCTTTGAACGTCAACCTGACCG 114326
QY 1261 ATGAGTGGCGCGTGTGCGCATGCGGCGGCGCATGCGCGGCGGACACGACATGAGCA 1320
DB 114325 ATGAGTGGCGCGTGTGCGCATGCGGCGGCGCATGCGCGGCGGACACGACATGAGCA 114266
QY 1321 GTCACTTGTGACACGACGTGCGCAGGGGTATGACCTTATCGACGACCTTCACACT 1380
DB 114265 GTCACTTGTGACACGACGTGCGCAGGGGTATGACCTTATCGACGACCTTCACACT 114206
QY 1381 CAACGCTGCGGCTAAATTTCTCGGTCTTGGGATCAAGAAATGCTTAATCAACCTTG 1440
DB 114205 CAACGCTGCGGCTAAATTTCTCGGTCTTGGGATCAAGAAATGCTTAATCAACCTTG 114146
QY 1441 CAATTTTGTGTCTTGAATCAACGCGCAGGTGCAAAAGGTTCATTAGGTCAATGAG 1500
DB 114145 CAATTTTGTGTCTTGAATCAACGCGCAGGTGCAAAAGGTTCATTAGGTCAATGAG 114086
QY 1501 TACTTAAATGAGTAAATATATCTCTGAT-TTTAAAGAGTCCACCATGAAATCACT 1559
DB 114085 TAAATTAATAGAGCAAAATCTTCTGATTAATAAGAGATCGACCATGAAATCACTATC 114026
QY 1560 ATCTGAAAGAGAGCAAGAGTGGGCAAGGATGCGCAAGTCTTAATGACACCTTGGCC 1619
DB 114025 ATCTGAAAGAGAGCAAGAGTGGGCAAGGATGCGCAAGGATGCGCAAGTCTTAATGAC 113966
QY 1620 AACAGGGTGAACCTTGGGGCTTGCAACAGGATCTCAGCAGTGAATCAACAGAG 1679

Db	113965	ACTAAGGGGCGGAACTTGGGGCTTCAACCTGATGTCACCTTTAGAGCACTTACCAAG	113906
Qy	1680	CTCATTCGATGATGAAAGCTGGGGAAGTGCATTCAGAACTGCAAGCATTTCTGTG	1739
Db	113905	CTCATTCGATGATGAAAGCTGGGGAAGTGCATTCAGAACTGCAAGCATTTCTGTG	113846
Qy	1740	GATGAAATCGTGGGACTAAACCCGTGACGATGAAAAAGCTACTTTTAAACCATTCGCAA	1799
Db	113845	GATGAAATCGTGGGACTAAACCCGTGACGATGAAAAAGCTACTTTTAAACCATTCGCAA	113786
Qy	1800	GAGTTCACGACCACTGACATCGTTGATGAAAGGCTTACAGCCGAGTGGTCAAC	1859
Db	113785	GAGTTCACGACCACTGACATCGTTGATGAAAGGCTTACAGCCGAGTGGTCAAC	113726
Qy	1860	CTGATTCATACGAAAGCTGACAGATGAGGCAAAAGATGCTGCAGATCCGTTGA	1919
Db	113725	CTGATTCATACGAAAGCTGACAGATGAGGCAAAAGATGCTGCAGATCCGTTGA	113666
Qy	1920	GTTCAATCTTGGCATCGGCGGAAACGG-CACATCGCTTTTCAATGAAACCATCATCTTCT	1978
Db	113665	GTTCAATCTTGGCATCGGCGGAAACGG-CACATCGCTTTTCAATGAAACCATCATCTTCT	113606
Qy	1979	CTGTACAGACGACAAAGCTCCAGGCGTGCACCTTAAACGTTGAGGACAAAGCTGCA	2038
Db	113605	CTGTACAGACGACAAAGCTCCAGGCGTGCACCTTAAACGTTGAGGACAAAGCTGCA	113546
Qy	2039	TTCTTCAACACCATGGAAGAGGTCCCAACCCAGCGCTCACCCAGGTTTGGGCACTTGG	2098
Db	113545	TTCTTCAACACCATGGAAGAGGTCCCAACCCAGCGCTCACCCAGGTTTGGGCACTTGG	113486
Qy	2099	TCCCGCGCGCAAAACATCGTGTGGTGGCACTGTGAAAGAAAGCCGACCTCCG	2158
Db	113485	TCCCGCGCGCAAAACATCGTGTGGTGGCACTGTGAAAGAAAGCCGACCTCCG	113426
Qy	2159	GGAACGTGGAAGGCGGACGTGCTGTTCCAGAGTTTCATCTGATGATGACAAAC	2218
Db	113425	GGAACGTGGAAGGCGGACGTGCTGTTCCAGAGTTTCATCTGATGATGACAAAC	113366
Qy	2219	ATGCCACATCATCGTTGATGAGGACGATCCAAAGCTGGAAGCGCTGATCATAC	2278
Db	113365	ATGCCACATCATCGTTGATGAGGACGATCCAAAGCTGGAAGCGCTGATCATAC	113306
Qy	2279	CGTTCATGGAACAATTAAAGCTGCGCTTGAAGAAAGAAAGAAAGTACTGTGGGGCTA	2338
Db	113305	CGTTCATGGAACAATTAAAGCTGCGCTTGAAGAAAGAAAGAAAGTACTGTGGGGCTA	113246
Qy	2339	TGCACACAGAACTTTCAGATTGGCGCTGCGTACATGTAATCTCTCCGACGGGACGGC	2398
Db	113245	TGCACACAGAACTTTCAGATTGGCGCTGCGTACATGTAATCTCTCCGACGGGACGGC	113186
Qy	2399	TCAATGATCCAAACGGAATGACGTGATGAGATACCTTCCAGCTTACCTTACACGACG	2458
Db	113185	TCAATGATCCAAACGGAATGACGTGATGAGATACCTTCCAGCTTACCTTACACGACG	113126
Qy	2459	ATCCAGGTTTCCCTTTCGACCAAGGCGACCGGCTGGGCTACACACCAAGCCGTTGA	2518
Db	113125	ATCCAGGTTTCCCTTTCGACCAAGGCGACCGGCTGGGCTACACACCAAGCCGTTGA	113066
Qy	2519	CCGACCGGAGGATGAGTGGAGGACGACCTGCGGACGCTCTTAAACCGGATGATCT	2578
Db	113065	CCGACCGGAGGATGAGTGGAGGACGACCTGCGGACGCTCTTAAACCGGATGATCT	113006
Qy	2579	ATGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2638
Db	113005	ATGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	112946
Qy	2639	TCTACACCGGCAACCTTAAATTTGAGGAAAGCGCGGCAACCCAAACCTTGTGAG	2698
Db	112945	TCTACACCGGCAACCTTAAATTTGAGGAAAGCGCGGCAACCCAAACCTTGTGAG	112886
Qy	2699	TGAGAGCCCAACTGGGCTGATGGGCGGATTCATCGCGCTTCCCTTAAATTCGCTTA	2758
Db	112885	TGAGAGCCCAACTGGGCTGATGGGCGGATTCATCGCGCTTCCCTTAAATTCGCTTA	112826
Qy	2759	TGACGGAACCCGCGGCTTTCACACCTTAAACCGGATCCCATGATGACCTTGAATG	2818
Db	112825	TGACGGAACCCGCGGCTTTCACACCTTAAACCGGATCCCATGATGACCTTGAATG	112766
Qy	2819	GATGATGTTGGAACAATGTTCTTGGGGCCCAAGCGGAAACCTCACCGGTCAGCGGTTG	2878
Db	112765	GATGATGTTGGAACAATGTTCTTGGGGCCCAAGCGGAAACCTCACCGGTCAGCGGTTG	112706
Qy	2879	TATACCGCTGACAGATCTTGAATACTGGGAATTCCTCGGTGAATCACTTTGACCTCA	2938
Db	112705	TATACCGCTGACAGATCTTGAATACTGGGAATTCCTCGGTGAATCACTTTGACCTCA	112646
Qy	2939	GATGATGCAACCTGTTCTGCTCTTATCTGTTCCGATGCTACATGTTGGGAATGCC	2998
Db	112645	GATGATGCAACCTGTTCTGCTCTTATCTGTTCCGATGCTACATGTTGGGAATGCC	112586
Qy	2999	CCAACCTTTTAAACCTTCCGAGTGAAGAACTGGCGAAGATCTGACGTCGATTTTCT	3058
Db	112585	CCAACCTTTTAAACCTTCCGAGTGAAGAACTGGCGAAGATCTGACGTCGATTTTCT	112526
Qy	3059	GTCACAAAGATGGAACCGAATCCACGATGAGTTACTCACTACGCAAGCTCTGACAGT	3118
Db	112525	GTCACAAAGATGGAACCGAATCCACGATGAGTTACTCACTACGCAAGCTCTGACAGT	112466
Qy	3119	GCGATATGTCGTGACAAAGCTTGAAGAACGACCTTCCGCTCTTGGAGAGATTGACG	3178
Db	112465	GCGATATGTCGTGACAAAGCTTGAAGAACGACCTTCCGCTCTTGGAGAGATTGACG	112406
Qy	3179	AGCTGATTTGGGCAATGAAATTTACGACACGAGTTGACGTAACCGGTTCTGATGCT	3238
Db	112405	AGCTGATTTGGGCAATGAAATTTACGACACGAGTTGACGTAACCGGTTCTGATGCT	112346
Qy	3239	GCTCTGCTGATGATGAGGCTGCGCGGACAGATGATCAACCAAGTTGACAGGAAG	3298
Db	112345	GCTCTGCTGATGATGAGGCTGCGCGGACAGATGATGATCAACCAAGTTGACAGGAAG	112286
Qy	3299	GATGGGTGACCTGCTGACCTGTCGCCGCAAGCTTTCATTTGGCAACCAAGCGATCTTAC	3358
Db	112285	GATGGGTGACCTGCTGACCTGTCGCCGCAAGCTTTCATTTGGCAACCAAGCGATCTTAC	112226
Qy	3359	AAGAGCTCTTCTCCCAAGGAGGAGTGGGAGTAACTGATCTGATTTAGGTTCTGAC	3418
Db	112225	AAGAGCTCTTCTCCCAAGGAGGAGTGGGAGTAACTGATCTGATTTAGGTTCTGAC	112166
Qy	3419	CTGTCCGATGACATCCGAGCAATATTTCTCGAGTGGATGCTGCTGCTTGTCTG	3478
Db	112165	CTGTCCGATGACATCCGAGCAATATTTCTCGAGTGGATGCTGCTGCTTGTCTG	112106
Qy	3479	TGATCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	3538
Db	112105	TGATCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	112046
Qy	3539	ACGATTAACAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	3598
Db	112045	ACGATTAACAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	111987
Qy	3599	GCTTCAAGGTGACATATTTGAAGATTAATTAATAAGGCTCTTGTGGCAATT	3658
Db	111986	GCTTCAAGGTGACATATTTGAAGATTAATTAATAAGGCTCTTGTGGCAATT	111927
Qy	3659	GTAACAATCTTTCGAAATCCCTTGAATGCAACAAATAACAGGTTTAATGTTGTTA	3718
Db	111926	GTAACAATCTTTCGAAATCCCTTGAATGCAACAAATAACAGGTTTAATGTTGTTA	111867
Qy	3719	GCTTTGAACAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	3778
Db	111866	GCTTTGAACAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	111807
Qy	3779	ATGACCATTAAGGACCTGCGCAACGATCTGCGGACATGAGGCGGAAAGCAACAT	3838
Db	111806	ATGACCATTAAGGACCTGCGCAACGATCTGCGGACATGAGGCGGAAAGCAACAT	111747


```

RESULT 5
E11760
LOCUS      E11760      6911 bp      DNA      linear      PAT 29-SEP-1997
DEFINITION Base sequence of sucrase gene.
ACCESSION  E11760
VERSION    E11760.1 GI:22025396
KEYWORDS   JP 1996196280-A/1.
SOURCE     Corynebacterium glutamicum
            Corynebacterium glutamicum
            Bacteria; Actinobacteriae; Actinobacteriales;
            Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE  1 [bases 1 to 6911]
            Sugimoto, M., Otona, K., Nagase, K., Tsuchiya, M., Matsui, Y.,
            Yoshihara, Y. and Nakamatsu, M.
            SUCROSE GENE ORIGINATED FROM CORYNEFORM BACTERIUM AND ITS
            PATENT: JP 1996196280-A 1 06-AUG-1996;
            AJINOMOTO CO INC
COMMENT     OS Brevibacterium lactofermentum
            PN JP 1996196280-A/1
            PD 06-AUG-1996
            PE 30-JAN-1995 JP 1995012361
            PI SUGIMOTO MASAKAZU, OTONA KIYOKO, NAGASE KAZUO, TSUCHIYA
            MAKOTO, PI MATSUI YUTAKA, YOSHIHARA YASUHIKO, NAKAMATSU WATARU PC
            C12N15/09, C12N1/21, C12P13/06//C12N9/24, (C12N1/21, C12R1.15); CC
            strandness: Double;
            CC topology: Linear;
            FH Key Location/Qualifiers
            FT source 1..6911
            FT /organism='Brevibacterium lactofermentum' FT
            FT /strain='ATCC13869'
            FT /plasmid='pSSM30BS'
            FT 342..1505
            FT /product='A protein having high homology to FT
            CDS /product='N-acetylglucosamine-6-phosphate deacetylase' FT
            FT 2338..3609
            FT /product='Sucrase'
            FT 4438..5358
            FT /product='A protein high homology to FT
            FT UDP-N-acetylmuramoyl
            FT alanyl-D-glutamyl-meso-6-diaminopimelate FT
            FT 5570..6577
            FT /product='A protein having homology to FT
            FT phospho-N-acetylmuramoyl pentapeptide FT
            FT 6577..
            FT /product='transferase'
            FT Location/Qualifiers
            FT 1..6911
            FT /organism='Corynebacterium glutamicum'
            FT /mol_type='genomic DNA'
            FT /db_xref='taxon:1718'
ORIGIN
Query Match 61.8%; Score 3687; DB 6; Length 6911;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 241 CAACAGCAGCGCTCAAGTCCGAGATTAATTAAGTATCCGTAGACATAAGACATCTA 300
DB 241 CAACAGCAGCGCTCAAGTCCGAGATTAATTAAGTATCCGTAGACATAAGACATCTA 300
QY 301 CGTCTATGCTTGTGCGAAGGAACCAATACTCAGAAAGATGGCAGAAATGTGATCATT 360
DB 301 CGTCTATGCTTGTGCGAAGGAACCAATACTCAGAAAGATGGCAGAAATGTGATCATT 360
QY 361 ATCAAGAAATGACAGGTCAACAGTAAATAATTAAGGAAAGATTTGTCCTCCCTCG 420
DB 361 ATCAAGAAATGACAGGTCAACAGTAAATAATTAAGGAAAGATTTGTCCTCCCTCG 420
QY 421 GGGTATTTGATGCGTTTCTCACTCGAAACCGCATCATACAGGAATCTCTCTGAGAC 480
DB 421 GGGTATTTGATGCGTTTCTCACTCGAAACCGCATCATACAGGAATCTCTCTGAGAC 480
QY 481 CAGCAGCTAAATAACGACAGATTTCCACCCGAACTCCACAGATTTGTCCTGATTTATG 540
DB 481 CAGCAGCTAAATAACGACAGATTTCCACCCGAACTCCACAGATTTGTCCTGATTTATG 540
QY 541 ATCTTCATATCAAGTGGAAACGATGGCGCTTTCTTACCGGAAACGACGACGCGCA 600
DB 541 ATCTTCATATCAAGTGGAAACGATGGCGCTTTCTTACCGGAAACGACGACGCGCA 600
QY 601 GGAACACCGCGCAGATATCAACGCGAACAATGGCAAGACCGATGTCACAGATGGTTT 660
DB 601 GGAACACCGCGCAGATATCAACGCGAACAATGGCAAGACCGATGTCACAGATGGTTT 660
QY 661 CGGCGCGCGCTGACGCACTGGCAGCGCAGGTGAAACCTTATTCCTTGTGTGAAGAG 720
DB 661 CGGCGCGCGCTGACGCACTGGCAGCGCAGGTGAAACCTTATTCCTTGTGTGAAGAG 720
QY 721 TCTGCTGTGGCGCATTAAGTCAAGAGGCGCTTTCTATCAAGCATGCGGTGTGTGCTC 780
DB 721 TCTGCTGTGGCGCATTAAGTCAAGAGGCGCTTTCTATCAAGCATGCGGTGTGTGCTC 780
QY 781 AAAACCCGAGATTTATTTTCCGCGCAACCAAGATCTTCCCGGGTGAATCCATGCGG 840
DB 781 AAAACCCGAGATTTATTTTCCGCGCAACCAAGATCTTCCCGGGTGAATCCATGCGG 840
QY 841 GAAAGGTTGATCAATATCATCAAGTACAGTGGCGCGAAATAGACATCTTTCTGAGCTTC 900
DB 841 GAAAGGTTGATCAATATCATCAAGTACAGTGGCGCGAAATAGACATCTTTCTGAGCTTC 900
QY 901 TCGATCTTGGCAGCGGACCAATCATCTTCTTCTTGGGCAACATGATCAATTTTG 960
DB 901 TCGATCTTGGCAGCGGACCAATCATCTTCTTCTTGGGCAACATGATCAATTTTG 960
QY 961 ATACCACTACAGGCGCAATTCCTTGGCTAAAGAAATAATGTACAGGCTACAGGCTACG 1020
DB 961 ATACCACTACAGGCGCAATTCCTTGGCTAAAGAAATAATGTACAGGCTACAGGCTACG 1020
QY 1021 ATTGTTCAATGCGATGCTCGCTGATCATATAGGCTCCCGGAGCGCTGAGCTTTGC 1080
DB 1021 ATTGTTCAATGCGATGCTCGCTGATCATATAGGCTCCCGGAGCGCTGAGCTTTGC 1080
QY 1081 TTGCTGCGGCGACGTCGCGGGAACGATATGTTGATGTTATGCGCGACGCGCTGATTTG 1140
DB 1081 TTGCTGCGGCGACGTCGCGGGAACGATATGTTGATGTTATGCGCGACGCGCTGATTTG 1140
QY 1141 CCGATGAAACGGTATGATCTGTTCCAAACAACGCTTTTATCATCAGGACCGCATAG 1200
DB 1141 CCGATGAAACGGTATGATCTGTTCCAAACAACGCTTTTATCATCAGGACCGCATAG 1200
QY 1201 AAGCGCGGAAATGCGACAGCGGTGATCATTTTGGCGTTTGAACGTCACGCTCACCG 1260
DB 1201 AAGCGCGGAAATGCGACAGCGGTGATCATTTTGGCGTTTGAACGTCACGCTCACCG 1260
QY 1261 ATGAGATGCGCGCTGCTGCGCATGCGCGCGCATTCGCGGGGACCAAGACATTAACGA 1320
DB 1261 ATGAGATGCGCGCTGCTGCGCATGCGCGCGCATTCGCGGGGACCAAGACATTAACGA 1320
QY 1321 GTCAAGTTGTGACCAAGCGTGGCAGGGGTATGACGTTATGACGCGACCTTCACACT 1380

```

Db 1321 GTGAGTTCGTGCAACGCTGCGAGGGGRTAGCGCTTATCGACGGACCTCTCAACCT 1380
 QY 1381 CAACGCTGCGCTAAATTTCTCGGTCTTGCGCATCAAGAAATCGTAAATCCAACTTG 1440
 Db 1381 CAACGCTGCGCTAAATTTCTCGGTCTTGCGCATCAAGAAATCGTAAATCCAACTTG 1440
 QY 1441 CAAATTTTGTGCTTTGACTGAAACGGCCAGGTGCAAAAAGTCCATTTAGTCAACAG 1500
 Db 1441 CAAATTTTGTGCTTTGACTGAAACGGCCAGGTGCAAAAAGTCCATTTAGTCAACAG 1500
 QY 1501 TACTTAAAGTACGAGTAAATCTATCTGATTTTAAAGAGTCCACCATATGAAATCACTA 1560
 Db 1501 TACTTAAAGTACGAGTAAATCTATCTGATTTTAAAGAGTCCACCATATGAAATCACTA 1560
 QY 1561 TCTGCAAGAAGAGCAAGAGTGGCAAAAGCAAGTTCAGTCTTAATCGACCTTGGCA 1620
 Db 1561 TCTGCAAGAAGAGCAAGAGTGGCAAAAGCAAGTTCAGTCTTAATCGACCTTGGCA 1620
 QY 1621 ACAAGGTGGAACCTTGGGGCTTGCAACAGATCTCTCAACATGATACCTACAGAGC 1680
 Db 1621 ACAAGGTGGAACCTTGGGGCTTGCAACAGATCTCTCAACATGATACCTACAGAGC 1680
 QY 1681 TCATTCGATGTATGAGCTGGGGAGTGTCACTTCAAGAACTGCAAGCATCTTGTG 1740
 Db 1681 TCATTCGATGTATGAGCTGGGGAGTGTCACTTCAAGAACTGCAAGCATCTTGTG 1740
 QY 1741 ATGAATAGCTGGAGCTAACCCGTGACGATGAAACAGCTACTTTAAACCATTCGCAAG 1800
 Db 1741 ATGAATAGCTGGAGCTAACCCGTGACGATGAAACAGCTACTTTAAACCATTCGCAAG 1800
 QY 1801 AGTTCACTGACCATCGACATGCTGTATGAGAGAGTCTACAGCCCAATGCTGCAAC 1860
 Db 1801 AGTTCACTGACCATCGACATGCTGTATGAGAGAGTCTACAGCCCAATGCTGCAAC 1860
 QY 1801 AGTTCACTGACCATCGACATGCTGTATGAGAGAGTCTACAGCCCAATGCTGCAAC 1860
 Db 1861 CTGATCCATACGAGAGCTGACAGATATGAGGCAAAAGATCGCTGCAAGTCCGTTGAAG 1920
 QY 1861 CTGATCCATACGAGAGCTGACAGATATGAGGCAAAAGATCGCTGCAAGTCCGTTGAAG 1920
 Db 1921 TTCAAAATCCTTGGCATCGGCGGAAGGCAATCGCTTTCACTTGAACATCATCTTCTGT 1980
 QY 1921 TTCAAAATCCTTGGCATCGGCGGAAGGCAATCGCTTTCACTTGAACATCATCTTCTGT 1980
 Db 1921 TTCAAAATCCTTGGCATCGGCGGAAGGCAATCGCTTTCACTTGAACATCATCTTCTGT 1980
 QY 1981 GTCAAGACTGACAAAGTCCAGGCGCTGCAACCTTAAACCTGTGAGAGAAACGCTTGAT 2040
 Db 1981 GTCAAGACTGACAAAGTCCAGGCGCTGCAACCTTAAACCTGTGAGAGAAACGCTTGAT 2040
 QY 2041 CTTCAACACCATGAGAGGTCCCAACCCACGCGCTACCCAGGTTTGGGCACTTTGTC 2100
 Db 2041 CTTCAACACCATGAGAGGTCCCAACCCACGCGCTACCCAGGTTTGGGCACTTTGTC 2100
 QY 2101 CCGCGGCAAAACATCGTGTGGTGCACTGGTGAAGGAAAGCCGACGCCATCCGCGG 2160
 Db 2101 CCGCGGCAAAACATCGTGTGGTGCACTGGTGAAGGAAAGCCGACGCCATCCGCGG 2160
 QY 2161 AACTGTGGAAGGCCAGTGTCTTCTTGGCCAGGTTCACTCTGTAGATGCACAACT 2220
 Db 2161 AACTGTGGAAGGCCAGTGTCTTCTTGGCCAGGTTCACTCTGTAGATGCACAACT 2220
 QY 2221 GGCACCATCATGCTTGGATGAGAGCAGATATCAAGCTGGAAGAGCTGATCACTACCG 2280
 Db 2221 GGCACCATCATGCTTGGATGAGAGCAGATATCAAGCTGGAAGAGCTGATCACTACCG 2280
 QY 2281 TCTCATGAGCAATTAAGCTGCTAGAAACAAAGAAAGAAAGTACTGTGTGGGCTATG 2340
 Db 2281 TCTCATGAGCAATTAAGCTGCTAGAAACAAAGAAAGAAAGTACTGTGTGGGCTATG 2340
 QY 2341 CACACGAACCTTCCAGTTTGGCGCTGCGTACCATGTGATCTCTCCGAGGGCAGGCTC 2400
 Db 2341 CACACGAACCTTCCAGTTTGGCGCTGCGTACCATGTGATCTCTCCGAGGGCAGGCTC 2400
 QY 2401 AATGATCCCAACGGAATGTACGTGATGAGATACCTTCAACGCTACTACCGACGAT 2460
 Db 2401 AATGATCCCAACGGAATGTACGTGATGAGATACCTTCAACGCTACTACCGACGAT 2460

Db 2401 AATGATCCCAACGGAATGTACGTGATGAGATACCTTCAACGCTACTACCGACGAT 2460
 QY 2461 CCAAGTTTCCCTTGGCACCAAGGCGACCGGCTGGGCTCACACCAACGCGCTTGACC 2520
 Db 2461 CCAAGTTTCCCTTGGCACCAAGGCGACCGGCTGGGCTCACACCAACGCGCTTGACC 2520
 QY 2521 GGAACGAGCGATTCAGATGAGAGCGACCTGGCCGAGCGCTTTTACCCGGAATCATCTAT 2580
 Db 2521 GGAACGAGCGATTCAGATGAGAGCGACCTGGCCGAGCGCTTTTACCCGGAATCATCTAT 2580
 QY 2581 GACCTGATGATGCTATTCGCGTGGAGCGATTTTACTAGAGGCACTTAACTTTTC 2640
 Db 2581 GACCTGATGATGCTATTCGCGTGGAGCGATTTTACTAGAGGCACTTAACTTTTC 2640
 QY 2641 TACACCGGCAACCTTAAATTTGACGAAAGCGCCCGCCACCAAACTTGTCCAAATC 2700
 Db 2641 TACACCGGCAACCTTAAATTTGACGAAAGCGCCCGCCACCAAACTTGTCCAAATC 2700
 QY 2701 GAGGACCCCACTGGGCTGATGGGCGGCAATTCATCGCCGTTGCGCTTAAATCGCTTATC 2760
 Db 2701 GAGGACCCCACTGGGCTGATGGGCGGCAATTCATCGCCGTTGCGCTTAAATCGCTTATC 2760
 QY 2761 GACGACCCGCGAGCGGTTTCAACCCCATTAACCGGATCCGATATCAGCCCTGATGCT 2820
 Db 2761 GACGACCCGCGAGCGGTTTCAACCCCATTAACCGGATCCGATATCAGCCCTGATGCT 2820
 QY 2821 GATGTTGGAACATGCTTCTTGGGCGCAACGCGAAACCTTCAACGCTGACGCTTCTA 2880
 Db 2821 GATGTTGGAACATGCTTCTTGGGCGCAACGCGAAACCTTCAACGCTGACGCTTCTA 2880
 QY 2881 TACCGCTGACGATCTTGAAGAACTGGGAAATCTCCGGGAAATGACCTTTGACCTGCT 2940
 Db 2881 TACCGCTGACGATCTTGAAGAACTGGGAAATCTCCGGGAAATGACCTTTGACCTGCT 2940
 QY 2941 GATGCAACAACTGTCTTCTCTCTGATCTCGTCCGATGCTATGATGGAATGCCCC 3000
 Db 2941 GATGCAACAACTGTCTTCTCTCTGATCTCGTCCGATGCTATGATGGAATGCCCC 3000
 QY 3001 AACCTTTTTCAGCTTCCGATGAGAAACCTGGCGAAGATCTGACGCTGCTGATTTTCTGT 3060
 Db 3001 AACCTTTTTCAGCTTCCGATGAGAAACCTGGCGAAGATCTGACGCTGCTGATTTTCTGT 3060
 QY 3061 CCAAGAGATTTGAGACCGAATCCACGATGAGGTTACTCACTAGCAAGCTCTACCAAGTGC 3120
 Db 3061 CCAAGAGATTTGAGACCGAATCCACGATGAGGTTACTCACTAGCAAGCTCTACCAAGTGC 3120
 QY 3121 GGAATATGCTGCAAGAGCTTGAAGAAAGCACTTCCGCGCTTGTGAGAGATTCAGCGAG 3180
 Db 3121 GGAATATGCTGCAAGAGCTTGAAGAAAGCACTTCCGCGCTTGTGAGAGATTCAGCGAG 3180
 QY 3181 TTGATTTTCCGATGAATTTTACGACCGCAGGTTGAGTAAACGCTTCTGATGCTGCG 3240
 Db 3181 TTGATTTTCCGATGAATTTTACGACCGCAGGTTGAGTAAACGCTTCTGATGCTGCG 3240
 QY 3241 CTGCTGGGCTGATGAGGCTGCGCGGAGAGTATGATCAACCCAAACGTTGCAAGAGAA 3300
 Db 3241 CTGCTGGGCTGATGAGGCTGCGCGGAGAGTATGATCAACCCAAACGTTGCAAGAGAA 3300
 QY 3301 TGGGTGCACTGCTGATCTGTGCCCCGCAAGCTTCACTTGGCGCAACCAACGCTTACCA 3360
 Db 3301 TGGGTGCACTGCTGATCTGTGCCCCGCAAGCTTCACTTGGCGCAACCAACGCTTACCA 3360
 QY 3361 GAGCTCTTCTTCCAGAGGAGAGTGGGAGGTTAATCAATCTGTAATTAGTTCTGAACCT 3420
 Db 3361 GAGCTCTTCTTCCAGAGGAGAGTGGGAGGTTAATCAATCTGTAATTAGTTCTGAACCT 3420
 QY 3421 GTCCAGATGACATCCGAGGCAATTTTCCCTCGAGTGGAGTGTCTGTTGCTG 3480
 Db 3421 GTCCAGATGACATCCGAGGCAATTTTCCCTCGAGTGGAGTGTCTGTTGCTG 3480
 QY 3481 GATCGTGAATGATGCTGCGCTAGCTGAGGTAACCTGGGCAATTTAGTATCCGAGAC 3540
 Db 3481 GATCGTGAATGATGCTGCGCTAGCTGAGGTAACCTGGGCAATTTAGTATCCGAGAC 3540

QY 3541 GATATATACAGCCATTGAGATTAATCTGAGGTATGACAGGTTTCTATGCTTTTCCGGCC 3600
DB 3541 GATATATACAGCCATTGAGATTAATCTGAGGTATGACAGGTTTCTATGCTTTTCCGGCC 3600
QY 3601 CTTCAGAGGTGACATTAATGAGATTAATGATTAATGAGGTCTTTTGGCGAATGTT 3660
DB 3601 CTTCAGAGGTGACATTAATGAGATTAATGATTAATGAGGTCTTTTGGCGAATGTT 3660
QY 3661 ACAAAATCTTGGCAAAATCCCTTGATC 3687
DB 3661 ACAAAATCTTGGCAAAATCCCTTGATC 3687
RESULT 6
126124 6911 bp DNA linear PAT 07-OCT-1996
LOCUS Sequence 4 from patent US 556776.
DEFINITION 126124
ACCESSION 126124
VERSION 1.1 GI:1605994
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6911)
AUTHORS Tsuchiya, M. and Miwa, K.
TITLE Sucrose gene derived from Corynebacterium
JOURNAL Patent: US 556776-A 4 17-SEP-1996;
FEATURES
source 1..6911
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 61.8%; Score 3687; DB 6; Length 6911;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGTCCGTGACGCCACCATTTGATGCTGCTGACAGGCTTCCGAGCTTTCTACATCT 60
DB 1 AGTCCGTGACGCCACCATTTGATGCTGCTGACAGGCTTCCGAGCTTTCTACATCT 60
QY 61 AGCTCCCGTCCGCGTGAAGGAGTCAATTAAGGAGTATCAGCCGCGTGAAGATTGGC 120
DB 61 AGCTCCCGTCCGCGTGAAGGAGTCAATTAAGGAGTATCAGCCGCGTGAAGATTGGC 120
QY 121 GAACCCATGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
DB 121 GAACCCATGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 181 TCTGCAAGTTTAAAGTATGATCAATCACTGGAAGGCTGAGTATTCAGTAGACCTG 240
DB 181 TCTGCAAGTTTAAAGTATGATCAATCACTGGAAGGCTGAGTATTCAGTAGACCTG 240
QY 241 CAACAGCAGGCTCAAGTCCGAAAGATTAATTAAGTATTAAGTATTAAGTATTAAGT 300
DB 241 CAACAGCAGGCTCAAGTCCGAAAGATTAATTAAGTATTAAGTATTAAGTATTAAGT 300
QY 301 CGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
DB 301 CGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 361 ATCAAGAAATGCAAGTCAAGGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 420
DB 361 ATCAAGAAATGCAAGTCAAGGATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 420
QY 421 GGGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 421 GGGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 481 CAGCAGCTTAAAGCAGGATTCACCCGAACTCCCAAGATTTGTTCCGGTTTATTTG 540
DB 481 CAGCAGCTTAAAGCAGGATTCACCCGAACTCCCAAGATTTGTTCCGGTTTATTTG 540

QY 541 ATCTTCAATATCAGGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
DB 541 ATCTTCAATATCAGGAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
QY 601 GGAACACCGCAGTATCAACCGCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 GGAACACCGCAGTATCAACCGCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 CGGCGCGGCTGACGACATGCGACGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB 661 CGGCGCGGCTGACGACATGCGACGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 TCTGCTGCTGCGGATTCACCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
DB 721 TCTGCTGCTGCGGATTCACCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
QY 781 AAAACCGGATTTGATTTTCCCGCAACCCCAAGATCTTCCCGGATGATCCATGCGG 840
DB 781 AAAACCGGATTTGATTTTCCCGCAACCCCAAGATCTTCCCGGATGATCCATGCGG 840
QY 841 GAAAGGTTGGATCAATCGATCAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
DB 841 GAAAGGTTGGATCAATCGATCAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
QY 901 TCGATCTGCGGAGCGGACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
DB 901 TCGATCTGCGGAGCGGACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 961 ATACCACTACGAGCGCAATTCCTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 961 ATACCACTACGAGCGCAATTCCTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 ATTGTTCAATGCGATGCTTCCGCTGATCAATGAGGCTCCGAGAGGAGGAGGAGGAGG 1080
DB 1021 ATTGTTCAATGCGATGCTTCCGCTGATCAATGAGGCTCCGAGAGGAGGAGGAGGAGG 1080
QY 1081 TTGCTGCGGACGTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
DB 1081 TTGCTGCGGACGTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
QY 1141 CGATGAAAGGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
DB 1141 CGATGAAAGGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
QY 1201 AAGCGCGGAGATGCGAGAGGAGTGAATTTGAGGAGTGAATTTGAGGAGTGAATTTGAG 1260
DB 1201 AAGCGCGGAGATGCGAGAGGAGTGAATTTGAGGAGTGAATTTGAGGAGTGAATTTGAG 1260
QY 1261 ATGAGTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 1261 ATGAGTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1321 GTCAAGTGTGACCAAGGAGGAGTGAATTTGAGGAGTGAATTTGAGGAGTGAATTTGAG 1380
DB 1321 GTCAAGTGTGACCAAGGAGGAGTGAATTTGAGGAGTGAATTTGAGGAGTGAATTTGAG 1380
QY 1381 CAACCGTGGCTGAAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
DB 1381 CAACCGTGGCTGAAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
QY 1441 CAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
DB 1441 CAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500
QY 1501 TACTTAAGTACGAGTAAATCTATCTGATTTTAAAGAGTCCCACTGAGAAATCACTA 1560
DB 1501 TACTTAAGTACGAGTAAATCTATCTGATTTTAAAGAGTCCCACTGAGAAATCACTA 1560
QY 1561 TCTGCAAG 1620
DB 1561 TCTGCAAG 1620

VERSION	BD165105.1	GI:27870917
KEYWORDS	JP 2002191370-A/2904.	
SOURCE	unidentified	
ORGANISM	unidentified	
REFERENCE	unclassified.	
AUTHORS	1 (bases 1 to 1983)	
TITLE	Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K., Yokoi,H., Tateishi,N., Senoo,A., Ikeda,M. and Ozaki,A.	
JOURNAL	Novel Polynucleotide Patent: JP 2002191370-A 2904 09-JUL-2002;	
COMMENT	KTOMA HAKKO KOGYO CO LTD OS Corynebacterium glutamicum PN JP 2002191370-A/2904 PD 09-JUL-2002 PF 15-DEC-2000 JP 2000405096 PI SATOISHI NAKAGAWA,HIROSHI MIZOGUCHI,SEIKO ANDO,MIKIO HAYASHI, PI KEIKO OCHIAI, PI HARUHIKO YOKOI,NAOKO TATEISHI,AKIHIRO SENOO,MASATO IKEYA,AKIO OZAKI PC C12N1/09,C12N15/09,C07K14/34,C07K16/12,C07K16/40,C12M1/00,PC C12N1/15, PC C12N1/19,C12N1/21,C12N5/10,C12N9/00,C12N9/02,C12P7/40,C12P13/PC 04,C12P13/08, PC C12P19/00,C12P19/34,C12P21/02,C12Q1/37,C12Q1/68,G01N33/53,PC G01N33/566, PC G01N33/569,G01N33/68,G01N37/00//C12P21/08,(C12N1/21,C12R1:15),PC (C12N1/21,C12R1:13), (C12N1/21,C12R1:01), (C12P13/08,C12R1:15), PC C12N15/00, PC C12N5/00,C12N15/00 CC Novel polynucleotide FH key FT source Location/Qualifiers 1..1983 /organism='Corynebacterium glutamicum'. Location/Qualifiers 1..1983 /organism='unidentified' /mol_type='genomic DNA' /db_xref='taxon:32644'	
ORIGIN	:	
Query Match	31.4%;	Score 1874.2; DB 6; Length 1983;
Best Local Similarity	96.6%;	Pred. NO. 0;
Matches 1915; Conservative	0;	Mismatches 68; Indels 0; Gaps 0;
OY	3779	ATGACACCTAAGGAACTTCGCCGAAGCATCTCTGCGCGCATTTGGCGGGAACAACACTT 3838
Pb	1	ATGACACCTAAGGAACTTCGCCGAAGCATCTCTGCGCGCATTTGGCGGGAACAACACTT 60
OY	3839	GTCGCGCGCGCACTGTGTGCAACGCGTTACGCGCTCGAGCTCAAGAACAACCAAGGATGTG 3898
Db	61	GTCGCGCGCGCACTGTGTGCAACGCGTTACGCGCTCGAGCTCAAGAACAACCAAGGATGTG 120
OY	3899	GATGCGCAAGCTCTGATGATGATCCAGATCTGAAAGGCACTTTGAACCTGGCGGCAATG 3958
Db	121	GATGCGCAAGCTCTGATGATGATCCAGATCTGAAAGGCACTTTGAACCGGTGGTATG 180
OY	3959	TTCCAGATCATCTGTGGGCGAGGCGATGGATCAATGTTTCAAGAACTGGATGACGCA 4018
Db	181	TTCCAGATCATCTGTGGGCGAGGCGATGGATCAATGTTTCAAGAACTGGATGACGCA 240
OY	4019	ACCTCCAAAGACATGCTGTGTCCAACAAGCAAGCTCAAGATGTTGTGGCTPACACGCC 4078
Db	241	ACCTCCAAAGACATGCTGTGTCCAACAAGCAAGCTCAAGATGTTGTGGCTPACACGCC 300
OY	4079	AACCTGTTCAAGCCGTCGTGTGAAAGTATTTGGCGGCAATTTTCGTCCCGCTGATTCATC 4138
Db	301	AACCTGTTCAAGCCGTCGTGTGAAAGTATTTGGCGGCAATTTTCGTCCCGCTGATTCATC 360
OY	4139	TTGGTTGGGCGGCTCTGCTCAAGCTATCAACATGTTGGTGGTGGCAGAGATCTGTC 4198
Db	361	TTGGTTGGGCGGCTCTGCTCAAGCTATCAACATGTTGGTGGTGGCAGAGATCTGTC 420
OY	4199	GATCCGCAATCACTGTGTGAAGATGTTCCCTCAATCAAGCAGGTGTTGCTGAAGATGATCAAC 4258

Db	421	GGTCGCGAATCACTGGTGGAGATGTTCCCTCAGATCAAGCGGTGTTGCTGAGATGATCAAC	480
Oy	4259	CTGATGGCATTGCGCCGCTTGCGCTTCTTGCAAGTGTGGTGGTTCAACCGCAACAAG	4318
Db	481	CTCATGGCACTGGCGCGTTGCGCCTTCTTGCAATGTGGTGGTTTCAACGCAACAAG	540
Oy	4319	CGTTCCGATGGCAATGAGTTCCTTGAGGCGCGCAATGTGTATGAGCATGAGTTCACAAC	4378
Db	541	CGTTCCGCGGCATGAGTTCCTGGGCGCGGATGTGTATGAGCATGAGTTCGAGC	600
Oy	4379	CTGGTTAACGGCTTACGACGTGGCCGCCAATCAATGACCGCGGCGAAATGCCAATGGTTC	4438
Db	601	TTGGTGAACGGCTTACGACGTGGCCGCCAATGAGCTCCGGCGAAATGCCAATGTGGTTC	660
Oy	4439	CTGGTTGAGTTGAGATGTGCTCAAGCTGGTTTACAGGAGCAACCGTCTCTGCTGGT	4498
Db	661	CTGTTTGGTTTGAATGTGCTCCAGCCGGTTTACAGGAGCAACCGTCTCTGCTGGT	720
Oy	4499	GTCCTTTGAGTTCTGGCAACGATGAGAGTTCTTGCAACAAGCGATCATAGGCACTGCA	4558
Db	721	GTTTCTTGAGATTCTGGCAACGATGAGAGTTCTTGCAACAAGCGATCATAGGCACTGCA	780
Oy	4559	GACTTCTGTATCAACCCAGTGTGAATCTGTGTGTCAACCGCTTCTTAAGTTATGCT	4618
Db	781	GACTTCTGTATCACTCAAGTGTGAATCTGTGTGTCAACCGATTCCTTAATATGATGCC	840
Oy	4619	ATTGATCCAGCAATGCGCTGGTGGTGAATCTGTGTCACACGGTCTGCAAGGACTAT	4678
Db	841	ATTGAGCCAGCAATGCGCTGGTGGTGAATGTGTGTCACACGGTCTTACAGGACTTAT	900
Oy	4679	GATTTCGGTGTCCAGTTCGGCGGTCTGTGCTTTGGTCTGTCTACTCAACCAATGGTATC	4738
Db	901	GATTTCGGTGTCCAGTTCGGCGGTCTGTGCTTTGGTCTGTCTACTCAACCAATGTATC	960
Oy	4739	ACTGATCGACCAAGTCCCTCCGSCCAATTGAGCTGAGGCTTCAACCAAGGATGATCC	4798
Db	961	ACTGATCGACCAAGTCCCTCCGSCCAATTGAGCTGAGGCTTCAACCAAGGATGATCC	1020
Oy	4799	TTCAATCTTCGCAACCGCATTCATGAGGCCAATATCGCGAGGGTGCAGCATGTTTGGCACTG	4858
Db	1021	TTCAATCTTCGCAACCGCATTCATGAGGCCAATATCGCCAGGGTGCAGCATGTTTGGCACTG	1080
Oy	4859	TTCTTCTTACGGAAGATGAAGCTGAAGGGCTTTCAGAGTGTCCGCGT	4918
Db	1081	TTCTTCTTACGGAAGATGAAGCTGAAGGGCTTTCAGAGTGTCCGCGT	1140
Oy	4919	GTTCTTGGTATTCAGAGGCTTGCATCTTCCGTGTGAACCTTCGCTGCGCTGCGCTTC	4978
Db	1141	GTTCTTGGTATTCAGAGGCTTGCATCTTCCGTGTGAACCTTCGCTGCGCTGCGCTTC	1200
Oy	4979	TACATTTGGTATCGGTATCCGACGCTATGCGTGGCGCTTGAATGACATTTGAATATAG	5038
Db	1201	TTCAATCGGTATCGGTATCCGACGCTATGCGTGGCGCTTGAATGACATTTGAATATAG	1260
Oy	5039	GCAATTCGATGGGCGCTGCAAGTTCCTTGAGTGTGTTCTATTTGATGCTCCAGATATG	5098
Db	1261	GCAATTCGATGGGCGCTGCAAGTTCCTTGAGTGTGTTCTATTTGATGCTCCAGATATG	1320
Oy	5099	GTCATGTTCTTGGTTGCGCGGTATGTTACTTTGTCAATCGCATTCGCGCGCATTTGCT	5158
Db	1321	GTCATGTTCTTGGTTGCGCGGTATGTTACTTTGTCAATCGCATTCGCGCGCATTTGCT	1380
Oy	5159	TATGAGCCTTTACTTGGTTCGCGCAAGGGACATGATCAAGATGCAACCGCGCTGCA	5218
Db	1381	TATGAGCCTTTACTTGGTTCGCGCAAGGGACATGATCAAGATGCAACCGCGCTGCA	1440
Oy	5219	GTCGCTCAGGAACGACCAAAAGCCGAGACAGAACCCGCGAGATTTTCAACGATATCC	5278
Db	1441	GTCGCTCAGGAACGACCAAAAGCCGAGACAGAACCCGCGAGATTTTCAACGATATCC	1500
Oy	5279	ACCATATCCAGGCACTTTGAACGGTGAAGCTATGCGACTGAGACGCTTACGAGATGCC	5338

Db 1501 ACCATCATCOAGCACCTTTGACCGGTGAAGCTATTGCACTGACGACGCTGACGATGCC 1560
 QY 5339 ATGTTTGGCAGCGGAAAGCTTGGCTCAGGTGTTGCGATGTTCCCAACCAAGGGCAGCTG 5398
 Db 1561 ATGTTTGGCAGCGGAAAGCTTGGCTCAGGTGTTGCGATGTTCCCAACCAAGGGCAGCTT 1620
 QY 5399 GTTTTCACGATGAGCGGAAAGATCGTGTGGCCCTTCCCATGTGTCACGCTTTGCGAGTC 5458
 Db 1621 GTTTTCACGATGAGCGGAAAGATCGTGTGGCCCTTCCCATGTGTCACGCTTTGCGAGTC 1680
 QY 5459 CGCAGTAAAGCTGAGGATGGTTCATATGATATCTTGAATGCACTTGGTTTCAGACCC 5518
 Db 1681 CGCAGTAAAGCTGAGGATGGTTCATATGATATCTTGAATGCACTTGGTTTCAGACCA 1740
 QY 5519 GTTAACTCAACGCGACGCACTTTAACCGGCTGAAGACAGGGCAGATGAATCAAGCA 5578
 Db 1741 GTTAACTCAACGCGACGCACTTTAACCGGCTGAAGACAGGGCAGATGAAGCAAGCA 1800
 QY 5579 GGGAGAGCTGCTGTGAATTCGATATGATATGACCACTTAAGGCTGACAGTTATGAGGTAC 5638
 Db 1801 GGGAGAGCTGCTGTGAATTCGATATGATATGACCACTTAAGGCTGACAGTTATGAGGTAC 1860
 QY 5639 ACGCCGATTTGTTTTCGATTAACAAGAAACCGGACTCTTAACAATTACGCTTTGGGC 5698
 Db 1861 ACGCCGATTTGTTTTCGATTAACAAGAAACCGGACTCTTAACAATTACGCTTTGGGC 1920
 QY 5699 GAAATTTGAAGCGGAGACCACTGCTCAACGTCGCAAGAAAGACGGTCCAGCAACA 5758
 Db 1921 GAAATTTGAAGCGGAGACCACTGCTCAACGTCGCAAGAAAGACGGTCCAGCAACA 1980
 QY 5759 CCA 5761
 Db 1981 CCA 1983
 RESULT 8
 AX122988 1983 bp DNA linear PART 11-MAY-2001
 LOCUS AX122988 Sequence 2904 from Patent EP1108790.
 DEFINITION AX122988
 VERSION AX122988.1 GI:14040476
 KEYWORDS
 SOURCE
 ORGANISM
 Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE
 AUTHORS
 1 Nakagawa, S., Mizoguchi, H., Ando, S., Hayaishi, M., Ochiai, K.,
 Yokoi, H., Tateishi, N., Senoh, A., Ikeda, M. and Ozaki, A.
 TITLE
 Novel polynucleotides
 Patent: EP 1108790-A 2904 20-JUN-2001;
 KYOMA HAKKO KOSYO CO., LTD. (JP)
 FEATURES
 source
 1. 1983
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 ORIGIN
 Query Match 31.4%; Score 1874.2; DB 6; Length 1983;
 Best Local Similarity 96.6%; Pred. No. 0;
 Matches 1915; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Db 121 GATCCGAAGAATCTGGATGATGATCAAGATCTGAAGGACGCTTTGAACGGGTGTATG 180
 QY 3959 TTCCAGATCATGTGGGCGCAGGCGATGTGATCATGTTTTCAAAGAACTCGATGACGA 4018
 Db 181 TTCCAGATCATGTGGGCGCAGGCGATGTGATCATGTTTTCAAAGAACTCGATGACGA 240
 QY 4019 AACTCCAAAGAAATGCGTGTGTCACAGAGCAGCTCAAGAGATGTTGTGCTAAACAGCC 4078
 Db 241 AACTCCAAAGAAATGCGTGTGTCACAGAGCAGCTCAAGAGATGTTGTGCTAAACAGCC 300
 QY 4079 AACTGTTTCAGCGCTGTGTGAAGGATTTGGCGGACATTTTTCGCGCTGATTCATATC 4138
 Db 301 AACTGTTTCAGCGCTGTGTGAAGGATTTGGCGGACATTTTTCGCGCTGATTCATATC 360
 QY 4139 TTGTTTGTGGGCGTGTGCTCATGTGCTCAACAAATGTGTTGTTGCGGAGATCTGTTTC 4198
 Db 361 TTGTTTGTGGGCGTGTGCTCATGTGCTCAACAAATGTGTTGTTGCGGAGATCTGTTTC 420
 QY 4199 GGTCCGCAATCATGCTGTGAGATGTTCCCTCAGATCAGCGGTTGCTGAGATGATCAAC 4258
 Db 421 GGTCCGCAATCATGCTGTGAGATGTTCCCTCAGATCAGCGGTTGCTGAGATGATCAAC 480
 QY 4259 CTGATGAGCATTCGCGCCGTTTCGCTTTCGCAAGTGTGTTGTTTCAACGCAACCAAG 4318
 Db 481 CTGATGAGCATTCGCGCCGTTTCGCTTTCGCAAGTGTGTTGTTTCAACGCAACCAAG 540
 QY 4319 CGTTTCGTTGCGCAATGAGTTCTCTGAGCGCGGCAATTGTTATGCGATGTTTCCCAACC 4378
 Db 541 CGTTTCGTTGCGCAATGAGTTCTCTGAGCGCGGCAATTGTTATGCGATGTTTCCCAACC 600
 QY 4379 CTGTTTAAACGCTAAGACGTGGCGCCGCAACGATGACCGGGGGGAAATGCAATGTGTTC 4438
 Db 601 TTGTGTAACGCTAAGACGTGGCGCCGCAACGATGACCGGGGGGAAATGCAATGTGTTC 660
 QY 4439 CTGTTTGTGTTGATGTTGCTCAAGCTGTTTACAGAGGACCGTCTTCTGTGCTGTG 4498
 Db 661 CTGTTTGTGTTGATGTTGCTCAAGCTGTTTACAGAGGACCGTCTTCTGTGCTGTG 720
 QY 4499 GTCTCTTGGATTTCTGGCAACGATCGAAGAGTTCTCTGACACAGCAGCTCATGGCAGTGA 4558
 Db 721 GTCTCTTGGATTTCTGGCAACGATCGAAGAGTTCTCTGACACAGCAGCTCATGGCAGTGA 780
 QY 4559 GACTTCCTGATACCCCAAGTGTGACTGCTGCTGCTCACCGGCTTCCCTTAAGTTTGTCT 4618
 Db 781 GACTTCCTGATACCCCAAGTGTGACTGCTGCTGCTCACCGGCTTCCCTTAAGTTTGTCT 840
 QY 4619 ATTGATCCAGCAATCGCTGAGTGTGATCTGTGTCAGACAGGTCGTGAGGAGACTATAT 4678
 Db 841 ATTGATCCAGCAATCGCTGAGTGTGATCTGTGTCAGACAGGTCGTGAGGAGACTATAT 900
 QY 4679 GATTTTGGTGTGTCAGTGTGGCGGTCTGTTTTCGCTGTGCTACTACCAATTCGTTATC 4738
 Db 901 GATTTTGGTGTGTCAGTGTGGCGGTCTGTTTTCGCTGTGCTACTACCAATTCGTTATC 960
 QY 4739 ACTGATTCGACACAGTCTTCCCGCAATGAGCTGAGAGCTGTTCAACCAAGAGTGGATTC 4798
 Db 961 ACTGATTCGACACAGTCTTCCCGCAATGAGCTGAGAGCTGTTTCAACCAAGAGTGGATTC 1020
 QY 4799 TTTCATCTTCGCAACCGCATTCATGCGCAATATCGCGAGGTCAGACATGTTTGGCAGTG 4858
 Db 1021 TTTCATCTTCGCAACCGCATTCATGCGCAATATCGCGAGGTCAGACATGTTTGGCAGTG 1080
 QY 4859 TTCTTCTTAGCGAAGATGAAGAAAGCTCAAGGGCTTTGACAGTGTTCAGGTCTTCGGCT 4918
 Db 1081 TTCTTCTTAGCGAAGATGAAGAAAGCTCAAGGGCTTTGACAGTGTTCAGGTCTTCGGCT 1140
 QY 4919 GTTCTTGGATTAACAGAGCTGCGATCTTCGCTGGAAGCTTCGCTGCGCTGGCGGCGTTTC 4978
 Db 1141 GTTCTTGGATTAACAGAGCTGCGATCTTCGCTGGAAGCTTCGCTGCGCTGGCGGCGTTTC 1200
 QY 4979 TACATTTGATTCGATACCGCAGCTATCGGTGCGCTTTGATTCGACTCTTTGATATCAAG 5038
 Db 1201 TACATTTGATTCGATACCGCAGCTATCGGTGCGCTTTGATTCGACTCTTTGATATCAAG 1260

QY 5039 GCAAGTTCGCTGGGCGCTGCAAGTTTCTTGGGTGTTGTTCTATGATGCTCCAGATATG 5098
 DB 1261 GCAGTGTGGTGGGCGCTGCAAGTTTCTTGGGTGTTGTTCTATGATGCTCCAGATATG 1320
 QY 5099 GTCATGTTCTTGGTGTGGCGGTAGTTACCTTGTGATCGATTCGAGCGGCGAGCATGCT 5158
 DB 1321 GTCATGTTCTTGGTGTGGCGGTAGTTACCTTGTGATTCGATTCGAGCGGCGAGCATGCT 1380
 QY 5159 TATGCGCTTTCATCTTGGTTCGCGCAACGCGAGCATTTGATTCAGATGCAACCGTGTCCA 5218
 DB 1381 TATGCGCTTTCATCTTGGTTCGCGCAACGCGAGCATTTGATTCAGATGCAACCGTGTCCA 1440
 QY 5219 GTGCTGCAAGAACCAACCAACCGAGAGAGACCCCGAGATTTTTCACATGCC 5278
 DB 1441 GTGCTGCAAGAACCAACCAACCGAGAGAGACCCCGAGATTTTTCACATGCC 1500
 QY 5279 ACCATCATCGAGGACCTTTGACCGGTGAGCTATCGACTGAGCAGCGTCAGCATGCC 5338
 DB 1501 ACCATCATCGAGGACCTTTGACCGGTGAGCTATCGACTGAGCAGCGTCAGCATGCC 1560
 QY 5339 ATGTTTCCGAGCGGAAAGCTTGGCTCAGGTGTCGATGTCGCCACCAAGGGGAGCTG 5398
 DB 1561 ATGTTTCCGAGCGGAAAGCTTGGCTCAGGTGTCGATGTCGCCACCAAGGGGAGCTTGA 1620
 QY 5399 GTTTTCCGAGCGGAAAGCTTGGCTCAGGTGTCGATGTCGCCACCAAGGGGAGCTG 5458
 DB 1621 GTTTTCCGAGCGGAAAGCTTGGCTCAGGTGTCGATGTCGCCACCAAGGGGAGCTTGA 1680
 QY 5459 CGCAGTAAAGCTGAGGATGTTCCATGATGATCTTGATGATGATGATGATGATGATGATG 5518
 DB 1681 CGCAGTAAAGCTGAGGATGTTCCATGATGATCTTGATGATGATGATGATGATGATGATG 1740
 QY 5519 GTTAACTTCAACGCGACGCACTTTAAACCGCTGAAAGAGCGGAGTGAATCAAAGCA 5578
 DB 1741 GTTAACTTCAACGCGACGCACTTTAAACCGCTGAAAGAGCGGAGTGAATCAAAGCA 1800
 QY 5579 GGGGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5638
 DB 1801 GGGGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
 QY 5639 AGCGGATGTTGTTTCAATTAACAAGAAACCGGACCTGTAACACTTACGGTTGGGC 5698
 DB 1861 AGCGGATGTTGTTTCAATTAACAAGAAACCGGACCTGTAACACTTACGGTTGGGC 1920
 QY 5699 GAAATTTGAAGCGGAGCCCACTGCTCAACGTCGCAAGAAAGAGCGGTGCACACACA 5758
 DB 1921 GAAATTTGAAGCGGAGCCCACTGCTCAACGTCGCAAGAAAGAGCGGTGCACACACA 1980
 QY 5759 CCA 5761
 DB 1981 CCA 1983

RESULT 9
 BD094213 1656 bp DNA linear PAT 27-AUG-2002
 LOCUS BD094213
 DEFINITION Genes for heat resistant enzymes of amino acid biosynthetic pathway
 derived from thermophilic coryneform bacteria.
 ACCESSION BD094213.1 GI:22639801
 VERSION WO 0125447-A/71.
 KEYWORDS Corynebacterium thermoaminogenes
 SOURCE Corynebacterium thermoaminogenes
 ORGANISM Corynebacterium thermoaminogenes
 Bacteria; Actinobacteria; Corynebacteriaceae; Corynebacterium.
 1 (bases 1 to 1656)
 HIRANO, S., NONAKA, G., MATSUZAKI, Y., AKIYOSHI, N., NAKAMURA, K.,
 KIMURA, E., OSUMI, T., MATSUI, K., KAWAHARA, Y., KURAHASHI, O.,
 NAKAMATSU, T. and SUGIMOTO, S.
 Genes for heat resistant enzymes of amino acid biosynthetic pathway
 derived from thermophilic coryneform bacteria
 Patent: WO 0125447-A 71 12-APR-2001;

COMMENT
 AJINOMOTO CO INC, SEIKO HIRANO, GEN NONAKA, YUMI MATSUZAKI, NAOKI
 AKIYOSHI, KANAE NAKAMURA, EIICHIRO KIMURA, TSUYOSHI OSUMI, KAZUHIKO
 MATSUI, YOSHIO KAWAHARA, OSAMU KURAHASHI, TSUYOSHI NAKAMATSU,
 SHINICHI SUGIMOTO
 OS Corynebacterium thermoaminogenes
 PN WO 0125447-A/71
 PD 12-APR-2001
 PF 04-OCT-2000 WO 2000JP06913
 PR 04-OCT-1999 JP 99P 282716, 01-NOV-1999 JP 99P 311147 PR
 PI SEIKO HIRANO, GEN NONAKA, YUMI MATSUZAKI, NAOKI AKIYOSHI, KANAE
 PI NAKAMURA,
 PI EIICHIRO KIMURA, TSUYOSHI OSUMI, KAZUHIKO
 MATSUI, YOSHIO KAWAHARA,
 PI OSAMU KURAHASHI, TSUYOSHI NAKAMATSU, SHINICHI SUGIMOTO PC
 C12N15/60, C12N15/54, C12N15/53, C12N15/31, C12N15/56, C12N9/88, PC
 C12N9/12,
 PC C12N9/04, C07K14/34, C12N9/26, C12N13/04
 CC
 FH Key Location/Qualifiers
 FT CDS (309)..(1595).
 FEATURES
 source Location/Qualifiers
 1..1656
 /organism="Corynebacterium thermoaminogenes"
 /mol_type="genomic DNA"
 /db_xref="taxon:161484"
 ORIGIN
 Query Match 24.4%; Score 1457; DB 6; Length 1656;
 Best Local Similarity 93.2%; Pred. No. 0;
 Matches 1545; Conservative 0; Mismatches 110; Indels 2; Gaps 2;
 QY 2031 ACGCTCAATCTTCAACACATCGAAGAGGCGCCACCAAGCCGTCACCCAGGGTTTGG 2090
 DB 1 ACGCCGATTTCTTCAACATATCAAGAGGTCCTCAACCAAGCGTTGACCAAGGCTTGG 60
 QY 2091 GCACCTTGTCCCGCGCGCAAAACATCGTGTGTGAGCACTGTGAAAGAAAGCCGAC- 2149
 DB 61 GTCATTTGTCCCGCGCGCAAAATATCGTGTGTGAGCACTGTGAAAGAAAGCA 120
 QY 2150 GCCATCGCGGAACTGTGAAAGGCCCACTGATCTCTTCTTCCAGGTTTCAATCTGTAG 2209
 DB 121 GCCATCGCGGAACTGTGAAAGGCCCACTGATCTCTTCTTCCAGGTTTCAATCTGTAG 180
 QY 2210 ATGCACACATGCCACCATCATCTGTTGATGAACAGAGATGATCAAGCTGAAACACGCT 2269
 DB 181 ATGCACACACGCGCACCATCATCTGTTGATGAACAGAGATGATCAAGCTGAAACACGCT 240
 QY 2270 GATCACTACCGTCTCATGAGCAATTAAAGCTGCGCTGAAGAAAGAAAGTACTGT 2329
 DB 241 GACCATTAACGTCATGAGCAATTAAAGCTGCGCTGAAGAAAGAAAGTACTGT 300
 QY 2330 GTGGGCGTATGACACACAACTTTCCAGTTTGGCCCTTGGCGTACCATGTGACTCTCCG 2389
 DB 301 GTGGGCGTATGACACACAACTTTCCAGTTTGGCCCTTGGCGTACCATGTGACTCTCCG 360
 QY 2390 AGGGAGGCTCAATGATCCCAAGCAAGTATGATGATGATGATGATGATGATGATGATGATG 2449
 DB 361 AGGGAGGCTCAATGATCCCAAGTATGATGATGATGATGATGATGATGATGATGATGATG 420
 QY 2450 ACCAGACGATTCAGAGTTTCCCTTTCGACCAAAAGCGACCTGGCTGGGCTCAACACCA 2509
 DB 421 ACCAGACGATTCAGAGTTTCCCTTTCGACCAAAAGCGACCTGGCTGGGCTCAACACCA 480
 QY 2510 CGCCGTTGACCGGACCGGAGGATGTCGATGAGAGCGACCTGGCGGAGCTTTTACCCGG 2569
 DB 481 CGCCGTTGACCGGACCGGAGGATGTCGATGAGAGCGACCTGGCGGAGCTTTTACCCGG 540
 QY 2570 ATGCATCTATGACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2629
 DB 541 ATGCATCTATGACCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
 QY 2630 TTAACCTTTCTACACCGGCAACTTAAATTAAGAGGAGCGCGCCACCCAAAC 2689

Db	601	TTAAACTTTTCTACACCGGCAACGAAAAATTGACGGCAAGCGCGCCACCAAAAC	660
Oy	2650	TTGTGGAAGTGGAGGACCCCACTGGGCTGATGGCGGCAATTGATGCGCTTAAAA	2749
Db	661	TCGTGGAAGTGGAGGACCCCACTGGGCTGATGGCGGCAATTGATGCGCTTAAAA	720
Oy	2750	ATCCGCTTATCGACGGGACCCGCGACGGTTTCAACCCCATTAACGGCATCCATGACA	2809
Db	721	ATCCGCTTATCGACGGGACCCGCGACGGTTTCAACCCCATTAACGGCATCCATGACA	780
Oy	2810	GCCCTGATGATGATGTTGGAAATGATGTTCTTGGGACCMAACGCAAAACCTCACCGTG	2865
Db	781	GCCCTGATGATGATGTTGGAAATGATGTTCTTGGGACCMAACGCAAAACCTCACCGTG	840
Oy	2870	CAGCGGTTCTATACCGCTCGACAGATCTTGGAAAATTCCTCCGCTGAAAACACT	2929
Db	841	CAGCGGTTCTATACCGCTCGACAGATCTTGGAAAATTCCTCCGCTGAAAACACT	900
Oy	2930	TTGACCTCAGATGATCACAACCTGGTTCTGCTCCTGATCTCGATCCGATGGCTACATG	2989
Db	901	TTGACCTCAGATGATCACAACCTGGTTCTGCTCCTGATCTCGATCCGATGGCTACATG	960
Oy	2990	GGGAAATGCCCAACTTTTATACGCTTCGCGATGAAGAAATCGGCAAGATCTCGATGC	3049
Db	961	GGGAATGCCCAACTTTTATACGCTTCGCGATGAAGAAATCGGCAAGATCTCGATGC	1020
Oy	3050	TGATTTTCGTTCACAAGGATTTGACCGAATCCAGATGAGGTTATCTACTACGCAACT	3109
Db	1021	TGATTTTCGTTCACAAGGATTTGACCGAATCCAGATGAGGTTATCTACTACGCAACT	1080
Oy	3110	CTGACCACTGCGGATATGTCGTGCACAAGCTTGAAGAAACGACTTCCGCGTTCGAG	3168
Db	1081	CTGACCACTGCGGATATGTCGTGCGCAAGCTTGAAGAAACGACTTCCGTCGTCGAG	1148
Oy	3170	GATTCAGGAGCTGGAATTTCCGCCATGAAATCTATAGCAACGAGGTTGCATTAACGTT	3229
Db	1141	GATTCAGGAGCTGGAATTTCCGTCATGAAATCTATAGCAACGAGGTTGCATTAACGTT	1200
Oy	3230	CTGATGCTCGGCTCGTGGGCTGAGATGGGAGCTGCCCCCGCAGAGTATGACCAACAGTTG	3289
Db	1201	CGAATGCTCGGCTTGTGGGCTGGAGTGGGATTTGCTGCAACGATGATCACCCAACAGTTG	1260
Oy	3280	CACAGGAAGATGGGTGCACTGCTGATCTGTCCTCCGCAAGCTTATTTGGCAACCAAG	3349
Db	1261	CGCAGGAAGATGGGTGCACTGCTGCAACGTTCTCGCAGGCTTCATTTGGTAACTCAATG	1320
Oy	3350	CGATTTACCAAGAGTCCCTTCCTCCAGAGGGGAGATCGGGGGTATTCAGATCTGATTAAG	3409
Db	1321	CGATTTACCAAGAGTCTTCTCTCCAGAGGGGAGATCGGGGGTATCTGATCTGATTAAG	1380
Oy	3410	GTTCTGAACTGTCCGAGTAAACATCCAGGCAATATTTCCCTCGAGTGGGATGGTGTCC	3469
Db	1381	GTTCTGAACTGTCCGAGTAAACATCCGACAAATGTTTCCCTCGAGTGGGATGGTGTCC	1440
Oy	3470	GTTTGTCTGTGAGATGCGATGCGATGCTGTCGCTGATGCTGAGGTAAACCTTGGCAATTAAG	3529
Db	1441	GTTTGTCTGTGAGATGCGATGCGATGCTGTCGCTGATGCTGAGGTAAACCTTGGCAATTAAG	1500
Oy	3530	TGATGCGCGAGCATATATACAGCAATGAGATTAACAGCAGATGAGACAGGTTTCATTGCG	3589
Db	1501	TGATGCGCGAGCATATATACAGCAATGAGATTAACAGCAGATGAGACAGGTTTCATTGCG	1560
Oy	3590	CTTTTCCGGGCTTCAAAAGTGACACTATTTGAGAGATTAAGTCATTAAGGGGCTTTTG	3649
Db	1561	C-TTTCGCCACCTTCAAAAGTGACACTATTTGAGAGATTAAGTCATTAAGGGGCTTTTG	1619
Oy	3650	TGGCGAATGTATACAATATCTTCGCAAAATCCCTTAT	3686
Db	1620	TGGCGAATGTATACAATATCTTCGCAAAATCCCTTAT	1656

LOCUS	AX069134	1527 bp	DNA	linear	PAT 25-JAN-2001
DEFINITION	Sequence 1 from Patent WO0102583.				
ACCESSION	AX069134				
VERSION	AX069134.1	GI:12579016			
KEYWORDS					
SOURCE					
ORGANISM	Corynebacterium glutamicum Corynebacterium glutamicum Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.				
REFERENCE	1 Pompejus, M., Kroegeer, B., Schroeder, H., Zelder, O. and Haberhauser, G. Orynebacterium glutamicum genes encoding phosphoenolpyruvat e: sugar phosphotransferase system proteins Patent: WO 0102583-A 1 11-JAN-2001; BASF AKTIENGESSELLSCHAFT (DE)				
FEATURES	Location/Qualifiers				
source	1..1527 /organism="Corynebacterium glutamicum" /mol_type="unassigned DNA" /db_xref="taxon:1718" 101..1507 /note="unnamed protein product; RKS00315" /codon_start=1 /transl_table=1 /protein_id="CAC27276.1" /db_xref="GI:12579017" /translation="MAWVPPSVLVNGVYVAATMAAGEMPWMSIFGLDVAOAGYOGTVLE VIVSVMIATIEKRLHKRLKGTADFLTTPULTLILNPLFLAIAGPMRWGVLIAHG LQGLVDRGFPVGGILFLGVLSPIYTLGHOSFPPIELBNROGGSFIFATSMANIAQ AGACLAIVFLAKSEKRLAGLAVSGVSLVLTETEPALTEGNRLKWPFFIGITAIAG ALILFNIKAVPALGAGPLGVSLIDAPDMWFLICAVVTEPIAAGAIAYGLYVRRR GSIDPDAFNAVPAAGTAKAEAPAEFNSDSTIIQADLTGEALILSSVSDMFAFGXKI GSGAIVPTKGOLVSPVSGKIVAFPSGHAFVARTKADGSDNVDIIMHIGDTVNLING THPEPLKKOGDEVKAGLELCEFDIDAIKAGYEVTTPILVSNYKKTGVNVTYGLGEIEI AGNLILNAVAKKENPATP"				
CDS					
ORIGIN					
Query Match	23.7%; Score 1412.6; DB 6; Length 1527;				
Best Local Similarity	95.7%; Pred. No. 0;				
Matches 1462; Conservative	0; Mismatches 64; Indels 1; Gaps 1;				
QY	4259 CTGATGGCATTTGGCGCCGCTTGGCGTTCTTGGCATATTTGGTTGGTTTACCCGCAACAAG 4318				
DB	1 CTGATGGCATTTGGCGCGCTTGGCGTTCTTGGCATATTTGGTTGGTTTACCCGCAACAAG 60				
OY	4319 CGTTTCGGTGGCAGTGAAGTTCCTGGGCGCGG-GCATTGGTATGAGCGATGCTCCCAAC 4377				
DB	61 CGTTTCGGCGGCATGAGTTCCTGGGCGCGCGCGCATTTGGTATGGCGATGCTGCCAG 120				
OY	4378 CTTGGTTAACGGCTACGACGTTGCGCCGCCACCATGACCGCGGGCAAAATGCCAATGTGGTC 4437				
DB	121 CTTGGTTAACGGCTACGACGTTGCGCGCCGCCACCATGCTGCGGGCGAAATGCAAAATGTGGTC 180				
OY	4438 CCGTCTTTGGTTGGATGTTGTGCTCAAGCGTGGTTACCAAGGCGACCGTGTCTCTGTGCTGGT 4497				
DB	181 CCGTCTTTGGTTGGATGTTGTGCTCAAGCGTGGTTACCAAGGCGACCGTGTCTCTGTGCTGGT 240				
OY	4498 GGTCTCTTGGATTTCTGGCAACGATCGAAGATTTCTGCACAAAGGCACTCATGGGCACTGC 4557				
DB	241 GGTTCCTTGGATTTCTGGCAACGATCGAAGATTTCTGCACAAAGGCACTCAAGGGCACTGC 300				
OY	4558 AGACTTCTCTGATCACCCAGTGTGACTCTGCTGCTCAACGGGCTTCCCTTAGGTTCAATTGC 4617				
DB	301 AGACTTCTCTGATCACTCCAAGTGTGAGAGTGTGCTCAACGGGATTCCTTAGATTCATTCGC 360				
OY	4618 TATTGGTTCAGCAATGGCGTGGGTGGGTGACTTGTCTGGACACACGGTCTGCAGGAGACTCTA 4677				
DB	361 CATTGGCGCGCAATGGCGTGGGTGGGTGGGTGCTGGACACACGGTCTGCAGGAGACTTTTA 420				
OY	4678 TGATTTGGGTGGTCCAAGTCGGCGGCTCTGCTTTTGGTCTGGTCTAATCAACCAATCGTAT 4737				
DB	421 TGATTTGGGTGGTCCAAGTCGGCGGCTCTGCTTTTGGTCTGGTCTAATCAACCAATCGCAT 480				


```

4738 CACTGCTGACACGAGTCTTCCGCAATTGAGCTGAGCTGTTCAACGAGGATGATC 4797
481 CACTGCTGACACGAGTCTTCCGCAATTGAGCTGAGCTGTTCAACGAGGATGATC 540
4798 CTTCACTTTCGCAACCGCATTCATGCGCAATATCCGCGAGGATGAGCATGTTGGCAT 4857
541 CTTCACTTTCGCAACCGCATTCATGCGCAATATCCGCGAGGATGAGCATGTTGGCAT 600
4858 GTTCTTCCAGGAGAGGAGAAAGCTAAGGAGGCTGAGGATGAGTCTCCG 4917
601 GTTCTTCCAGGAGAGGAGAAAGCTAAGGAGGCTGAGGATGAGTCTCCG 660
4918 TGTTCCTGATTAAGAGAGCTGCAATCTTGGTGAACCTTGGCTGGCTGGCT 4977
661 TGTTCCTGATTAAGAGAGCTGCAATCTTGGTGAACCTTGGCTGGCTGGCT 720
4978 CTACATGATGATGAGTACCGGAGCTATCGGTGGCTGATGATGATGATGATGAT 5037
721 CTACATGATGATGAGTACCGGAGCTATCGGTGGCTGATGATGATGATGATGAT 780
5038 GCGAGTGGCTGGGCGCTGAGAGTCTTGGGATGATGATGATGATGATGATGAT 5097
781 GCGAGTGGCTGGGCGCTGAGAGTCTTGGGATGATGATGATGATGATGATGAT 840
5098 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5157
841 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
5158 TTAATGCTCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5217
901 TTAATGCTCTTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
5218 AGTGCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5277
961 AGTGCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
5278 CACCATCATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5337
1021 CACCATCATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
5338 CATGTTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5397
1081 CATGTTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
5398 GATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5457
1141 GATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
5458 CCGCATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5517
1201 TCGCACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
5518 CTAATACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5577
1261 AGTAACCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
5578 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5637
1321 AGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
5638 CAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5697
1381 CAGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
5698 CGAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5757
1441 CGAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
5758 ACCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5784
1501 ACCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1527

```

```

RESULT 11
AX469820 1527 bp DNA linear PAT 09-AUG-2002
LOCUS Sequence 9 from Patent WO02051231.
DEFINITION
AX469820
VERSION
KEYWORDS
SOURCE
ORGANISM
Corynebacterium glutamicum
Corynebacterium glutamicum
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacterineae; Corynebacteriaceae; Corynebacterium.
REFERENCE
1
Pompejus, M., Kroege, B., Zelder, O. and Schroeder, H.
Genes of corynebacterium
Patent: WO 02051231-A 9 04-UTL-2002;
BASF AKTENGESSELSCRAFT (DE)
FEATURES
source
location/Qualifiers
1..1527
/organism="Corynebacterium glutamicum"
/mol_type="unassigned DNA"
/db_xref="taxon:1718"
101..1507
/notes="unnamed protein product; RKS00315"
/codon_start=1
/transl_table=1
/protein_id="CAD3524.1"
/db_xref="GI:22205072"
/translation="MAVVPSPVNGVDVAATMAAGEMPMSLFGDVAQXOGTVLP
VLVSWIATIEKHLKRLKGTADFLTPVLTLLTFLFIATIGPMKRVGVLHAG
LOGLYDGEFPGVGLFGLVYSPVIVTLOSFPETIELELNOGQSFIFAPASMANIQ
GACTAVFPLAKSEKLGKLAGSGLVSLVLTETEPALFVGLRLMPFIFIGDALIIG
ALITENIKAVLAGFLVNSIDADPMFLVCAVFEFAGALAVGLYLVRN
GSIDPNTAAPVPAGTAKARAPAEFNSSTIIQAPLTGEALIASVSDAMFSGTL
GSGVAIVTKQQLVSPVSKIVAPESGAPAPRTKACSDANVDIIMHIGPDYNLNG
THFPLKQGDVEVAGELCEFDIDAKAGYEVTTLVVSNTKTPVNTVGLGEIE
AGANILNVAKKEAVPATP"
ORIGIN
Query Match 23.7%; Score 1412.6; DB 6; Length 1527;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1462; Conservative 0; Mismatches 64; Indels 1; Gaps 1;
4259 CTGATGAGATCGGCGCGTTCGCTTGGCAGTGTGTTGATTCACCGCAACCAAG 4318
1 CTGATGAGATCGGCGCGTTCGCTTGGCAGTGTGTTGATTCACCGCAACCAAG 60
4319 CGTTTCGATGAGATGAGTTCCTGGGCGCG-CAATTGATGATGAGATGATGATGAT 4377
61 CGTTTCGATGAGATGAGTTCCTGGGCGCGCGCGATGATGATGATGATGATGATGAT 120
4378 CCGTGTAAAGGCTACGAGTGGCGCGCACCATGACCGCGGCGCAATGCAATGATGAT 4437
121 CTTGATGAAGCGCTACGAGTGGCGCGCACCATGATGATGATGATGATGATGATGAT 180
4438 CCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4497
181 CCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
4498 GGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4557
241 GGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
4558 AGACTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4617
301 AGACTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
4618 TATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4677
361 CATGAGCCAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
4678 TGAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4737

```

Db 421 TGATTCGGTGGTCCAGTCGGCGCTGCTCTTCGGTCTGTCTACTACCAATCGTCAT 480
 Qy 4738 CACTGGTTCGACCAAGTCTCTTCCCGCAATTGAGCTGGAGCTGTACACAGGCTGATC 4797
 Db 481 CACTGGTTCGACCAAGTCTCTTCCCGCAATTGAGCTGGAGCTGTATTAACAGGCTGATC 540
 Qy 4798 CTTCAATCTTCGCAACCGCATTCATGAGCCATATTCGGCAAGGGGTGAGCATGTTTGGCAT 4857
 Db 541 CTTCAATCTTCGCAACCGCATTCATGAGCCATATTCGGCAAGGGGTGAGCATGTTTGGCAT 600
 Qy 4858 GTTCTTCTAGCGAAGAGTGAAGAGCTCAAGGGCTTGGAGGTGCTTCAGTGTCTCCG 4917
 Db 601 GTTCTTCTAGCGAAGAGTGAAGAGCTCAAGGGCTTGGAGGTGCTTCAGTGTCTCCG 660
 Qy 4918 TGTTCCTGGTATTAACAGAGCTTCGCAATCTTCGGTGTGAACCTTCGGCTTCGGCTGCT 4977
 Db 661 TGTTCCTGGTATTAACAGAGCTTCGCAATCTTCGGTGTGAACCTTCGGCTTCGGCTGCT 720
 Qy 4978 CTACATTTGGTATCGATACCGCAGCTATCGGTGGCGCTTGGATTGCACTCTTGAATGCA 5037
 Db 721 CTTCAATCGGTATCGATACCGCAGCTATCGGTGGCGCTTGGATTGCACTCTTGAATGCA 780
 Qy 5038 GGCAGTGGCTTGGCGCTGACAGTCTTCGGGTGTGTTCTATGATGCTCCAGATAT 5097
 Db 781 GGCAGTGGCTTGGCGCTGACAGTCTTCGGGTGTGTTCTATGATGCTCCAGATAT 840
 Qy 5098 GGTCAATGTTCTTGGTGTGGCGGTAGTACCTTTGTGATCGCATTCGGCGACCGATTC 5157
 Db 841 GGTCAATGTTCTTGGTGTGGCGGTAGTACCTTTGTGATCGCATTCGGCGACCGATTC 900
 Qy 5158 TTATGGCCCTTTACTTGGTTCGCGCAAGCGGCAAGCATTCAGATGCAACCGCTGCTCC 5217
 Db 901 TTATGGCCCTTTACTTGGTTCGCGCAAGCGGCAAGCATTCAGATGCAACCGCTGCTCC 960
 Qy 5218 AGTGCCTGAGGAAAGCAACCAAGCGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5277
 Db 961 AGTGCCTGAGGAAAGCAACCAAGCGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1020
 Qy 5278 CACCATCATCCAGGCACTTTTGAACGGGTATGCACTGAGCAGCGTCAAGCATTC 5337
 Db 1021 CACCATCATCCAGGCACTTTTGAACGGGTATGCACTGAGCAGCGTCAAGCATTC 1080
 Qy 5338 CATGTTTGGCAGCGGAAAGCTTGGCTGAGTGTGAGTCGTCCTCCCAAGGGGAGCT 5397
 Db 1081 CATGTTTGGCAGCGGAAAGCTTGGCTGAGTGTGAGTCGTCCTCCCAAGGGGAGCT 1140
 Qy 5398 GGTTCACCAAGTGAAGCGAAGATCGTGTGCTTCCATCTGCTCAAGCTTTCGCAAT 5457
 Db 1141 AGTTCCTCGGTGAGTGAAGATCGTGTGCTTCCATCTGCTCAAGCTTTCGCAAT 1200
 Qy 5458 CCGCATTAAGCTGAGATGTTCCATGTTGATCTTGAATGCACTTGTGTCAGC 5517
 Db 1201 TCGCAACCAAGCTGAGATGTTCCATGTTGATCTTGAATGCACTTGTGTCAGC 1260
 Qy 5518 CGTAAACCTCAACGGGACGCACTTTAACCGGCTGAAGAGCAAGGCGATGAAGTCAAAAC 5577
 Db 1261 AGTAAACCTCAACGGGACGCACTTTAACCGGCTGAAGAGCAAGGCGATGAAGTCAAAAC 1320
 Qy 5578 AGGGAAGCTGTGTGATTCGATTAATGATGCACTTAAGGCTGCAAGTTATGAGTAAAC 5637
 Db 1321 AGGGAAGCTGTGTGATTCGATTAATGATGCACTTAAGGCTGCAAGTTATGAGTAAAC 1380
 Qy 5638 CAGGCGGATGTTGTTTGAATTAACAAGAAACCGGACCTGTAACACTTACGTTTGG 5697
 Db 1381 CAGGCGGATGTTGTTTGAATTAACAAGAAACCGGACCTGTAACACTTACGTTTGG 1440
 Qy 5698 CGAAATGAAGGGGAGCAACCTGCTCAAGCTGCAAGAAAGAAAGAGCGGTGCAAGCAAC 5757
 Db 1441 CGAAATGAAGGGGAGCAACCTGCTCAAGCTGCAAGAAAGAAAGAGCGGTGCAAGCAAC 1500
 Qy 5758 ACCATTAAGTGAACCTTGAAGTTCG 5784
 Db 1501 ACCATTAAGTGAACCTTGAAGTTCG 1527

RESULT 12
 AX065221 1342 bp DNA linear PAT 24-JAN-2001
 LOCUS
 DEFINITION
 Sequence 347 from Patent WO0100844.
 AX065221
 ACCESSION
 VERSION
 AX065221.1 GI:12542933
 KEYWORDS
 SOURCE
 ORGANISM
 Corynebacterium glutamicum
 Corynebacterium glutamicum
 Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 REFERENCE
 1 Pompeius, M., Kroege, B., Schroeder, H., Zelder, O. and Haberbauer, G.
 Corynebacterium glutamicum genes encoding proteins involved in
 carbon metabolism and energy production
 Patent: WO 0100844-A 347 04-JAN-2001;
 BASF AKTINSELSCHAFT (DE)
 JOURNAL
 FEATURES
 source
 1. 1342
 location/Qualifiers
 /organism="Corynebacterium glutamicum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:1718"
 101...1342
 /note="unassigned protein product; RXN00041"
 /codon_start=1
 /transl_table=1
 /protein_id="CAC25850.1"
 /db_xref="GI:12542934"
 /translation="MHTLSSLRPAVHYVTPPOGRINDPNMGVYVDGTLHYVYQHDGFP
 PPRKRTMAHTTTPLPORLQWHLDELALYDPDSYLDGYSAGAFPTBGTLLFY
 TGNLTKIDKRRATQMLVEDEPTGLMGSHHRSRPNPLIDGASFTHYRPMSPD
 DGKMWVIVGAQRENLTGAVALYRSTDLNWERSGITPDLSDAQGSAADLVPGGYM
 ECPNLTFRDEBTGEDLVLIJFCPGDLRIDEYTVHVASSDQGVVGVKLBETTRVL
 RGFSLDFGHEFYAPQVAVNGSDAVLVGMGLPADDDHPTVAREGMVHCLTVRKLHL
 RNHAIVQELLPEBESGVIRSVLISEPVRVIDIRNLSLEMGVRLSVRCGDRVAEV
 KPEGLVTDNDVTAIEITAGDQ"
 ORIGIN
 Query Match 22.1%; Score 1321.2; DB 6; Length 1342;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 1329; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 Qy 2238 ATGAAGCAGCATTCACAGTGAAGAAAGCTGATCATCGTCTCATGAGCAATTAA 2297
 Db 1 ATGAAGCAGCATTCACAGTGAAGAAAGCTGATCATCGTCTCATGAGCAATTAA 60
 Qy 2238 AGCTGCGCTAGAAAGAAAGAAAGTACTGTGTGGGCTATGCAACAGAACTTCCAG 2357
 Db 61 AGCTGCGCTAGAAAGAAAGAAAGTACTGTGTGGGCTATGCAACAGAACTTCCAG 120
 Qy 2358 TTTGGCCCTGGCTACATGTCATCTTCGCGAGGCGAGCTCAATGATCCCAAGCAAT 2417
 Db 121 TTTGGCCCTGGCTACATGTCATCTTCGCGAGGCGAGCTCAATGATCCCAAGCAAT 180
 Qy 2418 GTACGTGATGAGATACCTTCACGTTACTTACACAGACGATCCAGGTTTCCCTTGGC 2477
 Db 181 GTACGTGATGAGATACCTTCACGTTACTTACACAGACGATCCAGGTTTCCCTTGGC 240
 Qy 2478 ACCAAGGCAACCGGCTGGGCTACACAGACGCGGTGACCGGACCGGCGATTGCA 2537
 Db 241 ACCAAGGCAACCGGCTGGGCTACACAGACGCGGTGACCGGACCGGCGATTGCA 300
 Qy 2538 GTGACGACACTGCGCGAGCTCTTTACCCGAGTCACTATGACCTGATGATGCTGA 2597
 Db 301 GTGACGACACTGCGCGAGCTCTTTACCCGAGTCACTATGACCTGATGATGCTGA 360
 Qy 2598 TTCCGCTGAGCGGCTATTACTGACGCGACACTTAACTTTTCTACACCGGCAACTTAA 2657
 Db 361 TTCCGCTGAGCGGCTATTACTGACGCGACACTTAACTTTTCTACACCGGCAACTTAA 420
 Qy 2658 AATTGACGAAAGCGCGCGGCAACCAAACTTGTGCAAGTCAGAGACCCAACTGGGCT 2717

Db	482	TCAGCCCTGATGTGATGTGTTGGAAAAATGTCTTTGGGGCCCAACGGAAAACTCAACCG	541
Oy	2867	GTGCAGCGGTTCTATACCGCTTCGACAGATCTTGAATTAATCTGGGAATTCCTCCGGTGAATCA	2926
Db	542	GTGCAGCGGTTCTATACCGCTTCGACAGATCTTGAATTAATCTGGGAATTCCTCCGGTGAATCA	601
Oy	2927	CCTTTGACCTCAGTATGCAACAACCTGGTTCTGCTCTCGTATCTGTTCTCCGATGGCTACA	2986
Db	602	CCTTTGACCTCAGTATGCAACAACCTGGTTCTGCTCTCGTATCTGTTCTCCGATGGCTACA	661
Oy	2987	TGTGGGAATGCCCAACCTTTTATACGTTCCGGATGAAGAAACTGGCGAATCTCGACG	3046
Db	662	TGTGGGAATGCCCAACCTTTTATACGTTCCGGATGAAGAAACTGGCGAATCTCGACG	721
Oy	3047	TGCTGATTTTCTGTCCACAAGGATTTGGACCGAATCCAGATGAGTTTCTCACTACGCA	3106
Db	722	TGCTGATTTTCTGTCCACAAGGATTTGGACCGAATCCAGATGAGTTTCTCACTACGCA	781
Oy	3107	GCTCTGACCAAGTGGCGGATATGTCTGTGCAACAAGCTTGAAGAAACGACTTCGCGCTCTTGC	3166
Db	782	GCTCTGACCAAGTGGCGGATATGTCTGTGCAACAAGCTTGAAGAAACGACTTCGCGCTCTTGC	841
Oy	3167	GAGGATTCAGGAGCTGGATTTCCGGCCATGAAATTCATAGCAACCGCAGGTTTGCAGTAAACG	3226
Db	842	GAGGATTCAGGAGCTGGATTTCCGGCCATGAAATTCATAGCAACCGCAGGTTTGCAGTAAACG	901
Oy	3227	GTTCTGATGCTGGCTGCTGGGGCTTGGATGGGGGCTGCCCCGACAGGATGATCAACCAACAG	3286
Db	902	GTTCTGATGCTGGCTGCTGGGGCTTGGATGGGGGCTGCCCCGACAGGATGATCAACCAACAG	961
Oy	3287	TTGCACAGGAAGATGGGTGCACTGCTGCACTGTGCCCCGCAAGCTTCAATTTGGCGAAC	3346
Db	962	TTGCACAGGAAGATGGGTGCACTGCTGCACTGTGCCCCGCAAGCTTCAATTTGGCGAAC	1021
Oy	3347	ACCGGATTTACCAAGAGTCTCTCTTCCCAAGGGGGAATCGGGGGTATTCAGATCTGTAT	3406
Db	1022	ACCGGATTTACCAAGAGTCTCTCTTCCCAAGGGGGAATCGGGGGTATTCAGATCTGTAT	1081
Oy	3407	TAGGTTCTGAACCTGTCCGAGTAGACATCCGAGCAATATTTCCCTCGAGTGGGATGGTG	3466
Db	1082	TAGGTTCTGAACCTGTCCGAGTAGACATCCGAGCAATATTTCCCTCGAGTGGGATGGTG	1141
Oy	3467	TCCGTTTGTCTGTGATCGTATGCTGTCGTCCGTAAGCTGAGGTAAACCTGCGCAT	3526
Db	1142	TCCGTTTGTCTGTGATCGTATGCTGTCGTCCGTAAGCTGAGGTAAACCTGCGCAT	1201
Oy	3527	TATGTATTCGCGACATATATACAGCCATTGAGATATCTGCAAGGTGATGACAGTTTCAT	3586
Db	1202	TATGTATTCGCGACATATATACAGCCATTGAGATATCTGCAAGGTGATGACAGGTTCAT	1261
Oy	3587	TCCGTTTCCGGGCGCTTCAAAGGTGACCTATTGAGACA	3625
Db	1262	TCCGCTTTCGCGGCTTTCAAAGGTGACCTATTGAGACA	1299
RESULT 14			
LOCUS	AX122989	1299 bp	DNA linear PAT 11-MAY-2001
DEFINITION	Sequence 2905 from Patent EP1108790.		
ACCESSION	AX122989		
VERSION	AX122989.1	GI:14040477	
KEYWORDS			
SOURCE			
ORGANISM	Corynebacterium glutamicum		
REFERENCE	Corynebacterium glutamicum		
AUTHORS	Bacteria; Actinobacteriota; Actinobacteridae; Actinomycetales;		
	Corynebacteriineae; Corynebacteriaceae; Corynebacterium.		
	1 Nakagawa,S., Mizoguchi,H., Ando,S., Hayashi,M., Ochiai,K.,		
	Yokoi,H., Tateishi,N., Senoh,A., Ikeda,M. and Ozaki,A.		
TITLE	Novel polynucleotides		
JOURNAL	Patent: EP 1108790-A 2905 20-JUN-2001;		
	KYOMA HAKKO KOGYO CO., LTD. (JP)		

FEATURES	Location/Qualifiers
source	1..1299
	/organism="Corynebacterium glutamicum"
	/mol_type="unassigned DNA"
	/db_xref="taxon:1718"
ORIGIN	
Query Match	21.2%; Score 1267.8; DB 6; Length 1299;
Best Local Similarity	99.0%; Pred. No. 0;
Matches 1286; Conservative	0; Mismatches 12; Indels 1; Gaps 1;
Oy	2327 TGTGTGGGGCGATACACAGAACTTTCCAGTTTGGGCCCTGGCTACATGTGACTCTC 2386
Db	2 TGTGTGGGGCGATACACAGAACTTTCCAGTTTGGGCCCTGGCTACATGTGACTCTC 61
Oy	2387 CGCAGGGCAGGCGCTCAATGATCCCAACGGAAATGATCGTGTGATGAGATACCTCCACGCT 2446
Db	62 CGCAGGGCAGGCGCTCAATGATCCCAACGGAAATGATCGTGTGAGATACCTCCACGCT 121
Oy	2447 ACTACCAAGCAGATCCAGGTTTCCCTTCGACCAAAAGCGACCGGCTGGCTCAACCA 2506
Db	122 ACTACCAAGCAGATCCAGGTTTCCCTTCGACCAAAAGCGACCGGCTGGCTCAACCA 181
Oy	2507 CCAAGCGGCTGACCGGACCGGACGGATTTGCAATGTGAAGGACCTGCGGACGCTTTTACC 2566
Db	182 CCAAGCGGCTGACCGGACCGGACGGATTTGCAATGTGAAGGACCTGCGGACGCTTTTACC 241
Oy	2567 CGAATGCATCCTATGACTGGATGTGATTCGCGGTGAGCGGTATTTTACTGACGGCA 2626
Db	242 CGAATGCATCCTATGACTGGATGTGATTCGCGGTGAGCGGTATTTTACTGACGGCA 301
Oy	2627 CACTTAAACTTTTCTACACCGGCACCTAAATAATGACGGAAGCGCGCGCACCCAAA 2686
Db	302 CACTTAAACTTTTCTACACCGGCACCTAAATAATGACGGAAGCGCGCGCACCCAAA 361
Oy	2687 ACCTTGTGAAAGTGAAGACCCCAACTGGGCTGATGGGCGGCAATTCATGCGCTTGGCTTA 2746
Db	362 ACCTTGTGAAAGTGAAGACCCCAACTGGGCTGATGGGCGGCAATTCATGCGCTTGGCTTA 421
Oy	2747 AAAATCCGCGTTATCGAGCGACCCGCGCAGCGGTTTCAACCCATTACCGCGATCCCATGA 2806
Db	422 AAAATCCGCGTTATCGAGCGACCCGCGCAGCGGTTTCAACCCATTACCGCGATCCCATGA 481
Oy	2807 TCAGCCCTGATGTGATGTGTGGAACAATGTTCTTGGGGGCCCAACCGCAAAACTCTACA 2866
Db	482 TCAGCCCTGATGTGATGTGTGGAACAATGTTCTTGGGGGCCCAACCGCAAAACTCTACA 541
Oy	2867 GTGCAGCGGTTCTATACCGCTCGACAGATCTTGAAAACTGGGAATTCCTCGGTTAAATCA 2926
Db	542 GTGCAGCGGTTCTATACCGCTCGACAGATCTTGAAAACTGGGAATTCCTCGGTTAAATCA 601
Oy	2927 CCTTGAACCTCAGGATGACCAACCTGGTTCTGCTCCTGAATCTCGTCCGATGGCATCA 2986
Db	602 CCTTGAACCTCAGGATGACCAACCTGGTTCTGCTCCTGAATCTCGTCCGATGGCATCA 661
Oy	2987 TGTGGGAATGCCCCAACCTTTTACGCTTGCGCATGAAGAAACTGCGAAGATCTCGACG 3046
Db	662 TGTGGGAATGCCCCAACCTTTTACGCTTGCGCATGAAGAAACTGCGAAGATCTCGACG 721
Oy	3047 TGTGATTTTTCTGTCCAAAGATTGGAACCGAATCCACGATGAGTTACTCACTAGSCAA 3106
Db	722 TGTGATTTTTCTGTCCAAAGATTGGAACCGAATCCACGATGAGTTACTCACTAGSCAA 781
Oy	3107 GCTTGCACCAATGGGGAATATGTGCTGCACAAAGCTTGAAGAAACGACTTCGCGCTTTGC 3166
Db	782 GCTTGCACCAATGGGGAATATGTGCTGCAGCAAGCTTGAAGAAACGACTTCGCGCTTTGC 841
Oy	3167 GAGGATTCAGGAGCTGGATTTTGGCGCATGAAATCTTACGCAACCGCAGGTTGCAATAACG 3226
Db	842 GAGGATTCAGGAGCTGGATTTTGGCGCATGAAATCTTACGCAACCGCAGGTTGCAATAACG 901
Oy	3227 GTTGTGATGCTGTGGCTGTGTGGGCTGTGATGGGGCTGCCGCGCAGATGATTCACCAACAG 3286


```

Db      1141 AACCGTGGCCGCCAAATTCGGAAGTACGATCAGAAATGTTAAATCCAAACCTGT 1200
Qy      1442 AAATTTTGGTCTTTGACTCAACGGCCAGGTGCAAAAGTTCATTAGTCACTCACT 1501
Db      1201 AAATTTTGGTCTTTGACTCAACGGCCAGTTCACAGTTCATTAGTCACTCACT 1260
Qy      1502 ACTTAAAGTACGAGTAAACTATCTG 1528
Db      1261 AATTAAATACGAGCAAACTTCTG 1287

```

Search completed: March 10, 2005, 03:48:13
 Job time : 24435 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 13:35:43 ; Search time 2740 Seconds
(without alignments)
12895.962 Million cell updates/sec

Title: US-10-019-284B-1

Perfect score: 5969
Sequence: 1 agtcgcgtgcagccacatc.....tgatatcttcgcgcgtgaa 5969

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq.16dec04:*
1: _geneseqn1980s:*
2: _geneseqn1990s:*
3: _geneseqn2000s:*
4: _geneseqn2001as:*
5: _geneseqn2001bs:*
6: _geneseqn2002as:*
7: _geneseqn2002bs:*
8: _geneseqn2003as:*
9: _geneseqn2003bs:*
10: _geneseqn2003cs:*
11: _geneseqn2003ds:*
12: _geneseqn2004as:*
13: _geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5969	100.0	5969	4	AAf32543 Brevibact
2	5656.2	94.8	349980	5	AAh68533 C glutam
3	3687	61.8	6911	2	AAf34177 Corynefor
4	1874.2	31.4	1983	5	AAh67869 C glutam
5	1457	24.4	1656	5	AAf87497 Corynebact
6	1412.6	23.7	1527	5	AAf81528 C glutam
7	1412.6	23.7	1527	6	ABh65346 DNA encod
8	1321.2	22.1	1342	4	AAf71533 Corynebact
9	1267.8	21.2	1299	5	AAh67870 C glutam
10	1237.4	20.7	1287	4	AAf71567 Corynebact
11	1237.4	20.7	1287	4	AAf71567 Corynebact
12	1110.4	18.6	1152	5	AAh67872 C glutam
13	1049.8	17.6	1109	5	AAf31529 Corynebact
14	762	12.8	882	4	AAf71571 Corynebact
15	671.8	11.3	759	5	AAh67871 C glutam
16	474.2	7.9	479	5	AAf87433 Corynebact
17	466.6	7.8	498	5	AAf87432 Corynebact
18	322.8	5.4	110000	10	ADf77343_05 Continuation (6 of
19	294	4.9	1543	10	ADD13201 C. glutam
20	232.2	3.9	2913	4	AAh54100 S. epider

ALIGNMENTS

C	21	215.2	3.6	1059	4	AAf71769 Corynebact
C	22	215.2	3.6	1059	4	AAh56088 C. glutam
C	23	208	3.5	777	8	ACA30091 Prokaryot
C	24	192.6	3.2	1212	8	ACA30089 Prokaryot
C	25	192	3.2	975	6	ABK73538 Bacillus
C	26	183.2	3.1	1452	11	ABD00072 Klebsiell
C	27	181.4	3.0	1386	11	ABD00428 Streptoc
C	28	181	3.0	1491	10	ADH84532 Enterococ
C	29	177.6	3.0	1334	8	ACF73163 Staphyloc
C	30	177.6	3.0	30246	2	AAV74367 Staphyloc
C	31	176.2	3.0	861	8	ACF74550 Staphyloc
C	32	174.2	2.9	1380	6	ABK73394 Bacillus
C	33	173.8	2.9	2127	8	ABK42306 Streptoco
C	34	173.8	2.9	110000	10	ABh55454_16 Continuation (17 o
C	35	172.2	2.9	1965	10	ABX07668 S. pneumo
C	36	172.2	2.9	3695	2	AAV52334 Streptoco
C	37	168.8	2.8	906	13	ADS02113 Staphyloc
C	38	167.8	2.8	1932	13	ADR92012 Novel S.
C	39	163.6	2.7	1884	4	AAh55732 Streptoco
C	40	162.2	2.7	1428	11	ACH96011 Klebsiell
C	41	161	2.7	1380	5	AAh67868 C glutam
C	42	157	2.6	11330	4	AAh59544 Propionib
C	43	157	2.6	11330	8	ACF64473 Propionib
C	44	154.2	2.6	9769	2	AAV52163 Streptoco
C	45	154.2	2.6	110000	10	ABh55454_15 Continuation (16 o

RESULT 1	AAf32543	standard; DNA; 5969 BP.
ID	AAf32543	
XX	AAf32543;	
AC	XX	
DT	11-SEP-2003	(revised)
DT	20-APR-2001	(first entry)
DE	Brevibacterium lactofermentum	sucrose PTS enzyme II DNA SEQ ID NO:1.
XX	XX	
KW	Brevibacterium lactofermentum;	sucrose PTS enzyme II; sucrose;
KW	phosphoenolpyruvate:carboxylate	phosphotransferase system; glucose;
KW	coryneform bacterium;	phosphoenolpyruvate-sugar transport system; ds.
XX	XX	
OS	Corynebacterium	glutamicum.
XX	XX	
FH	Key	Location/Qualifiers
FT	CDS	3779..5764
FT		/*tag= a
FT		/product= "sucrose PTS enzyme II"
PN	MO200102584-A1.	
PD	11-JAN-2001.	
XX	XX	
PF	30-JUN-2000;	2000MO-JP004348.
XX	XX	
PR	02-JUL-1999;	99JP-00189512.
XX	XX	
PA	(AJIN) AJINOMOTO CO INC.	
XX	XX	
PI	Izui M, Sugimoto M, Nakamatsu T, Kurahashi O;	
XX	XX	
DR	WPI: 2001-138150/14.	
XX	XX	
PT	P-Psdb; AAB69080.	
XX	XX	
PT	Gene encoding sucrose phosphoenolpyruvate-sugar transport system enzyme	
PT	II obtained by cassette ligation-mediated amplification of downstream	
PT	domain of coryneform bacterium sucrose gene, with sucrose-binding	
XX	XX	
PS	Claim 3; Page 22-29; 45pp; Japanese.	

1921 TTCAAACTCTGGCATCGGCGGAAAAGGCAATCGCTTTCATTGAACCATCATCTTCTCT 1980
1921 TTCAAACTCTGGCATCGGCGGAAAAGGCAATCGCTTTCATTGAACCATCATCTTCTCT 1980
1981 GTCAGAGCTGACAAAGGTCCAGGCGCTGCACCTTAAACTGTGAGAGACAAACCGCTGAT 2040
1981 GTCAGAGCTGACAAAGGTCCAGGCGCTGCACCTTAAACTGTGAGAGACAAACCGCTGAT 2040
2041 CTTGAAACACCATGAAAGAGTCCCAACCCAGCCGTCACCCAGGTTTGGGCACTTTGTC 2100
2041 CTTGAAACACCATGAAAGAGTCCCAACCCAGCCGTCACCCAGGTTTGGGCACTTTGTC 2100
2101 CCGCGGCGGCAAAATCATGTGTGTGGCACTGTGTAAGGAAAAGCGGAGCCATCCGCGG 2160
2101 CCGCGGCGGCAAAATCATGTGTGTGGCACTGTGTAAGGAAAAGCGGAGCCATCCGCGG 2160
2161 AACTGTGAAGGCGCCAGTGAAGTCTTGGCCAGGTTCCATCTGTAGTGCACAACAT 2220
2161 AACTGTGAAGGCGCCAGTGAAGTCTTGGCCAGGTTCCATCTGTAGTGCACAACAT 2220
2221 GCCACCATCATCGTTGATGAGAGACAGCATATCAAGCTGGAAGAAAGCTGATCACTACG 2280
2221 GCCACCATCATCGTTGATGAGAGACAGCATATCAAGCTGGAAGAAAGCTGATCACTACG 2280
2281 TCTCATGAGAGCAATTAAAGCTGCGCTAGAAACAAAGAAAGTACTGTGTGGGCTATG 2340
2281 TCTCATGAGAGCAATTAAAGCTGCGCTAGAAACAAAGAAAGTACTGTGTGGGCTATG 2340
2341 CACACAGAACTTCCAGTTTGCGCCCTGCGTACATGTGACTCTCCGAGAGGACAGCTC 2400
2341 CACACAGAACTTCCAGTTTGCGCCCTGCGTACATGTGACTCTCCGAGAGGACAGCTC 2400
2401 AATGATCCAAAGGAAATGTAAGTGTGATGAGATACCTTCAAGCTTACTTACAGACAGAT 2460
2401 AATGATCCAAAGGAAATGTAAGTGTGATGAGATACCTTCAAGCTTACTTACAGACAGAT 2460
2461 CAGAGTTTCCCTTGGCAGCAAAAGGCGACCGGCTGGGCTCAACAGCAACAGCCGTTGAC 2520
2461 CAGAGTTTCCCTTGGCAGCAAAAGGCGACCGGCTGGGCTCAACAGCAACAGCCGTTGAC 2520
2521 GGAACCGCAAGATGAGAGTGAAGAGCACTGCGCGAGCTCTTTACCCGATGATCTTAT 2580
2521 GGAACCGCAAGATGAGAGTGAAGAGCACTGCGCGAGCTCTTTACCCGATGATCTTAT 2580
2581 GAACCTGATGATGCTATTCCGGTGAAGCGGATTTTACTGACGCGCACTTAACTTTTC 2640
2581 GAACCTGATGATGCTATTCCGGTGAAGCGGATTTTACTGACGCGCACTTAACTTTTC 2640
2641 TACACCGGCAACCTTAAATTTGACGAAAGCGCGGCGCAACCACTTGTGCAAGTC 2700
2641 TACACCGGCAACCTTAAATTTGACGAAAGCGCGGCGCAACCACTTGTGCAAGTC 2700
2701 GAGGACCCCAATCGGCTGATGAGGCGGCACTTCACTGCGCTTAAATTCGCTTATC 2760
2701 GAGGACCCCAATCGGCTGATGAGGCGGCACTTCACTGCGCTTAAATTCGCTTATC 2760
2761 GACGAGCCGCGCAGCGGTTTCAACCCATTACCGGATCCCATGATCAGCCCTGATGAT 2820
2761 GACGAGCCGCGCAGCGGTTTCAACCCATTACCGGATCCCATGATCAGCCCTGATGAT 2820
2821 GATGTGTGAACAATGTTCTTGGGGGCCAAGCGGAAACCTTCAACGCGTGCAGCGTTCTA 2880
2821 GATGTGTGAACAATGTTCTTGGGGGCCAAGCGGAAACCTTCAACGCGTGCAGCGTTCTA 2880
2881 TACCGCTGACAGATCTTGAATACTGGGAATTTCTCCGCTGAATCACTTTGACCTCAGT 2940
2881 TACCGCTGACAGATCTTGAATACTGGGAATTTCTCCGCTGAATCACTTTGACCTCAGT 2940
2941 GATGCAACAACCTGTGCTCTCTCTGATCTGTTCCGATGGCTACATGTGGGAATGCCCC 3000
2941 GATGCAACAACCTGTGCTCTCTCTGATCTGTTCCGATGGCTACATGTGGGAATGCCCC 3000

3001 AACCTTTTACGCTTCGCGATGAGAAGAACTGGGGAAGATCTCGACGCTGATTTTCTGT 3060
3001 AACCTTTTACGCTTCGCGATGAGAAGAACTGGGGAAGATCTCGACGCTGATTTTCTGT 3060
3061 CCACAGAGATGAGAGCAATCCACAGATGAGTTACTACTACGCAAGCTCTGACCAAGTC 3120
3061 CCACAGAGATGAGAGCAATCCACAGATGAGTTACTACTACGCAAGCTCTGACCAAGTC 3120
3121 GGAATGTGTGACACAGTTGAAGAACAGACCTTCCGCTCTTGGAGAGATTCAAGCAG 3180
3121 GGAATGTGTGACACAGTTGAAGAACAGACCTTCCGCTCTTGGAGAGATTCAAGCAG 3180
3181 CTGGATTTGGGCGATGAATTTACAGACCGGAGTTGAGTAAAGGTTCTGATGCTGAG 3240
3181 CTGGATTTGGGCGATGAATTTACAGACCGGAGTTGAGTAAAGGTTCTGATGCTGAG 3240
3241 CTTGTGGGCTGATGGGCTGCGCGGAGATGATCAACCAAGTTGACAGAGAAAG 3300
3241 CTTGTGGGCTGATGGGCTGCGCGGAGATGATCAACCAAGTTGACAGAGAAAG 3300
3301 TGGGTGCACTGCTGACTGTGCTGCGCGCAAGCTTCAATTTGCGCAACGCGATCTAC 3360
3301 TGGGTGCACTGCTGACTGTGCTGCGCGCAAGCTTCAATTTGCGCAACGCGATCTAC 3360
3361 GAGCTCCTTCTCCAGAGGAGGAGTCCGCGGTAATCAGATCTGTATAGCTTGAACCT 3420
3361 GAGCTCCTTCTCCAGAGGAGGAGTCCGCGGTAATCAGATCTGTATAGCTTGAACCT 3420
3421 GTCCGAGTGAACATCCGAGGCAATATTTCCCTCGAGTGGAGTGTCCGTTGCTGAG 3480
3421 GTCCGAGTGAACATCCGAGGCAATATTTCCCTCGAGTGGAGTGTCCGTTGCTGAG 3480
3481 GATCGTATGATGATGCTGCGGTGACCTGAGGTAAGGTAACCTGCGGAATTAATGAT 3540
3481 GATCGTATGATGATGCTGCGGTGACCTGAGGTAAGGTAACCTGCGGAATTAATGAT 3540
3541 GATTAATACGCAATGAGTAATCTGAGGTGATGAGAGGTTCAATTCGCTTCCGCG 3600
3541 GATTAATACGCAATGAGTAATCTGAGGTGATGAGAGGTTCAATTCGCTTCCGCG 3600
3601 CTTCAAGGTGACATATTTGAGAGATTAAGTCAATTAAGGCTCTTTTGGCGAATGT 3660
3601 CTTCAAGGTGACATATTTGAGAGATTAAGTCAATTAAGGCTCTTTTGGCGAATGT 3660
3661 ACAAAATCTTGGCAAAATCCCTTGAATCGGACCAAAATAACGTTTAATATTTTAC 3720
3661 ACAAAATCTTGGCAAAATCCCTTGAATCGGACCAAAATAACGTTTAATATTTTAC 3720
3721 TTTTGAACAAATCATGATGATGATTTTGTGTTCTTCCGCTTAAGAGAAATTCAT 3780
3721 TTTTGAACAAATCATGATGATGATTTTGTGTTCTTCCGCTTAAGAGAAATTCAT 3780
3781 GGAACATTAAGAGCTCGGCGCAACGCAATCTGCGGACATTTGGCGGAGAACATTTGT 3840
3781 GGAACATTAAGAGCTCGGCGCAACGCAATCTGCGGACATTTGGCGGAGAACATTTGT 3840
3841 CGCGCGCGCACTGTGCAACGCGTTTACGCTGCTGCTCAAAAGCACCAAGATGTGA 3900
3841 CGCGCGCGCACTGTGCAACGCGTTTACGCTGCTGCTCAAAAGCACCAAGATGTGA 3900
3901 TCGCCAAAGTCTGATGATGATCCAGATCTGAAAGGCACTTTGAAACCTGGCGCATGTT 3960
3901 TCGCCAAAGTCTGATGATGATCCAGATCTGAAAGGCACTTTGAAACCTGGCGCATGTT 3960
3961 CAGATCAATCGTGGGCGAGGAGTGGATCATGTTTCAAAAGAACTGATGACGCAAC 4020
3961 CAGATCAATCGTGGGCGAGGAGTGGATCATGTTTCAAAAGAACTGATGACGCAAC 4020
4021 CTTCAAGACATCGTGTGTCTCAGAGACAGTCAAGATGTTGTGCTTAAACAGCCAA 4080
4021 CTTCAAGACATCGTGTGTCTCAGAGACAGTCAAGATGTTGTGCTTAAACAGCCAA 4080
4081 CTGTTCAAGCGGTGCTGTGAAGTATTTGGCGGCAATTTTCCGCTGATCCATCTT 4140
4081 CTGTTCAAGCGGTGCTGTGAAGTATTTGGCGGCAATTTTCCGCTGATCCATCTT 4140

XX 18-DEC-2000; 2000EP-00127688.
XX
XX 16-DEC-1999; 99JP-00377484.
XX
XX 07-APR-2000; 2000JP-00159162.
XX
XX 03-AUG-2000; 2000JP-00280988.
XX
XX (KYOWA) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX WPI; 2001-376931/40.
XX
XX Novel polynucleotides derived from *Corynebacterium* bacteria, for identifying
XX PT mutation point of a gene, measuring expression of a gene, analyzing
XX expression profile or pattern of a gene and identifying homologous gene.
XX
XX PS Disclosure; SEQ ID NO 7068; 246bp + Sequence Listing; English.
XX
XX The present invention provides a number of nucleotide and protein
XX sequences from the *Corynebacterium* *Corynebacterium glutamicum*. These
XX are useful for identifying the mutation point of a gene derived from a
XX mutant of *Corynebacterium* *Corynebacterium*, measuring expression amount and analysing
XX the expression profile or expression pattern of a gene derived from
XX *Corynebacterium* *Corynebacterium*, and identifying a homologue of a gene derived from
XX *Corynebacterium* *Corynebacterium*. *Corynebacterium* bacteria are useful for producing amino
XX acids, nucleic acids, vitamins, saccharides and organic acids,
XX particularly L-lysine. The present sequence is a nucleic acid described
XX in the exemplification of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from the European Patent Office
XX
XX SQ Sequence 349980 BP; 81250 A; 97718 C; 90621 G; 80391 T; 0 U; 0 Other;

Query Match 94.8%; Score 5656.2; DB 5; Length 349980;
Best Local Similarity 97.1%; Pred. No. 0;
Matches 5793; Conservative 0; Mismatches 173; Indels 3; Gaps 3;

QY 1 AGTCGCTGACGCGCACCATTTGATGTGTGTCACGAGCTTGGCGAGGCTTTTACATCT 60
DB 115585 AGTCGCTGACGCGCACCATTTGATGTGTGTCACGAGCTTGGCGAGGCTTTTACATCT 115526
QY 61 ACCCTCCCGTGGCGGTGAGTGGGTCATTAACGGGTGGATCAACGCCGGTGAAGTTGG 120
DB 115525 ACACCTCCCGTGGCGGTGAGTGGGTCATTAACGGGTGGATCAACGCCGGTGAAGTTGG 115466
QY 121 GAACCATGTGTCTCTGTGGGTGAGGGAAGAGTGGGGGAGAGTTTTCAGTG 180
DB 115465 GAACCATGTGTCTCTGTGGGTGAGGGAAGAGTGGGGGAGAGTTTTCAGTG 115406
QY 181 TCTGCAAGTTTAAAGTTATGATCATCAGCTTGAAGGCTGAGGTATTCAGTGA 240
DB 115405 TCTGCAAGTTTAAAGTTATGATCATCAGCTTGAAGGCTGAGGTATTCAGTGA 115346
QY 241 CAAACAGAGGCTTCAAGTCCGAAGATTAATTAACCTTGATCCGTGACATTAAGCATTA 300
DB 115345 CAAACAGAGGCTTCAAGTCCGAAGATTAATTAACCTTGATCCGTGACATTAAGCATTA 115286
QY 301 CGTCCATAGCTTCTGGAAGGAACCAATTAACCTCAAGAAAGAGGAGAGTGTGCATT 360
DB 115285 CGTCCATAGCTTCTGGAAGGAACCAATTAACCTCAAGAAAGAGGAGAGTGTGCATT 115226
QY 361 ATCAAGAAATGAGGTCAAGAGTTAAATAATGAGGAGAGATTTTCCCTCCCTCG 420
DB 115225 ATCAAGAAATGAGGTCAAGAGTTAAATAATGAGGAGAGATTTTCCCTCCCTCG 115166
QY 421 GGGTGAATGATGCTTTCTCAACTGAAAACGGCATCATCAGGAATCTTCTGGAAG 480
DB 115165 GGGTGAATGATGCTTTCTCAACTGAAAACGGCATCATCAGGAATCTTCTGGAAG 115106
QY 481 CAGACCTAAAGAGGAGATTCACCCGCAACCTCCAGATGTTCCCGGTTTATTG 540
DB 115106 CAGACCTAAAGAGGAGATTCACCCGCAACCTCCAGATGTTCCCGGTTTATTG 113966

DB 115105 CAGACCTAAAGAGGAGATTCACCCGCAACCTCCAGATGTTCCCGGTTTATTG 115045
QY 541 ATCTTCAATATACCGGTGAAAACGGTGGCGCTTCTTACAGGAAACGAGACCGAGCGA 600
DB 115045 ATCTTCAATATACCGGTGAAAACGGTGGCGCTTCTTACAGGAAACGAGACCGAGCGA 114986
QY 601 GGAACACCGCGGATGATACCGCGGAACATGAGGAGAGTGTGCAACATATGTTT 660
DB 114985 GGAATGCGCGCGATATACCGCGGAACATGAGGAGAGTGTGCAACATATGTTT 114926
QY 661 CGGCGCGCGCTGACCGCATGCGAGCGAGGTGAAAACCTTATTCCTTGTGGAAGAG 720
DB 114925 CGGCGCGCGCTGACCGCATGCGAGCGAGGTGAAAACCTTATTCCTTGTGGAAGAG 114866
QY 721 TCTGCTGTGCGGCAATTCATCTGAGGGCCCTTTTATCAACGATCCGCTTGTGCTC 780
DB 114865 GCTGCTGTGCGGCAATTCATCTGAGGGCCCTTTTATCAACGATCCGCTTGTGCTC 114806
QY 781 AAAACCCGGAATTTCTTTTCCGGCAACCAAGATCTTCCGGGTGATCCATGCGG 840
DB 114805 AAAACCCGGAATTTCTTTTCCGGCAACCAAGATCTTCCGGGTGATCCATGCGG 114746
QY 841 GAAAAGTTGATCAATTCATCAGATGACCGCGAAACCTGACATCTTTTGAAGCTTC 900
DB 114745 GAAAAGTTGATCAATTCATCAGATGACCGCGAAACCTGACATCTTTTGAAGCTTC 114686
QY 901 TCGATCTGCGGCAAGCGCACCATCATCTTCTTCTGCGGACACTGATGAGATTTTG 960
DB 114685 TCGATCTGCGGCAAGCGCACCATCATCTTCTTCTGCGGACACTGATGAGATTTTG 114626
QY 961 ATACCATACGAGCGCAATGCTTGGCTAAAGAAAAATGACCGGTCACGCTACG 1020
DB 114625 ATACCATACGAGCGCAATGCTTGGCTAAAGAAAAATGACCGGTCACGCTACG 114566
QY 1021 ATTTGTTCAATGATGCTTCCGCTGATCATATGAGGCTCCCGAGCGGTGCGCTTTC 1080
DB 114565 ATTTGTTCAATGATGCTTCCGCTGATCATATGAGGCTCCCGAGCGGTGCGCTTTC 114506
QY 1081 TTGCTGCGGCAAGTGGCGGAGACGATATGTTGATGATGATGATGATGATGATGAT 1140
DB 114505 TTGCTGCGGCAAGTGGCGGAGACGATATGTTGATGATGATGATGATGATGATGAT 114446
QY 1141 CCGATGGAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 114445 CCGATGGAAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 114386
QY 1201 AAGCGCGCGAATGACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGAT 1260
DB 114385 AAGCGCGCGAATGACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGAT 114326
QY 1261 ATGAGTGGCGCGTGTGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 114325 ATGAGTGGCGCGTGTGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 114266
QY 1321 GTCAAGTGGTGAACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
DB 114265 GTCAAGTGGTGAACAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 114206
QY 1381 CAACGCTGCGCGCTAAATTTCTGCTTGGGATCAAGAAATGCTAAATCAACCTCG 1440
DB 114205 CAACGCTGCGCGCTAAATTTCTGCTTGGGATCAAGAAATGCTAAATCAACCTCG 114146
QY 1441 CAAATTTGTGCTTTTGAATCAACGCGTCAAGTCAAAAGTCCATTTAGTCAAG 1500
DB 114145 CAAATTTGTGCTTTTGAATCAACGCGTCAAGTCAAAAGTCCATTTAGTCAAG 114086
QY 1501 TACTTAAATGAAGTAAATCAATCTGAT - TTTAAAGAGTCCCAACATGAAATCACT 1559
DB 114085 TACTTAAATGAAGTAAATCAATCTGAT - TTTAAAGAGTCCCAACATGAAATCACT 114026
QY 1560 ATCTGCAAGAGAGAGAGATGCGCAAGAGTTCAGTCTAATGCACTTCCTGCGC 1619
DB 114025 ATCTGCAAGAGAGAGAGATGCGCAAGAGTTCAGTCTAATGCACTTCCTGCGC 113966

QY 1620 AACAGGGTGAACCTTGGGCTTGGCAAGAGATCTCAACAAGTACCTAACCAAG 1679
Db 113965 ACTAAGGGGGAGACCTTGGGGCTTGGCACTGATGCTCACTTTTGGAGCACTTACCAAG 113906
QY 1680 CTCATTGCGATGATGAAGCTGGGGAAGTGTCACTTCAAGACCTGCAAGGCACTTCTGTG 1739
Db 113905 CTCATTGCGATGATGAAGCTGGGGAAGTGTCACTTCAAGACCTGCAAGGCACTTCTGTG 113846
QY 1740 GATGAATACGTGGGACTAAACCCGTGACGATGAAAAACAGTACTTTAAAAACATTGCAAA 1799
Db 113845 GATGAATACGTGGGACTTTAAACGGCGACAGTGAACAGCTACTTCAAAACATTGTA 113786
QY 1800 GAGTTCACCTGACCAACATGCAATCTGTTGAAGAAGTCTTACAGCCCAAGTGTGCAAC 1859
Db 113785 GAGTTCACCTGACCAACATGCAATCTGTTGAAGAAGTCTTACAGCCCAAGTGTGCAAC 113726
QY 1860 CCTGATCCATAAGAGCAGCTGACAGATGAGGCAAGATGCTGCAAGATCCGTTGA 1919
Db 113725 CCTGATCCATAAGAGCAGCTGACAGATGAGGCAAGATGCTGCAAGATCCGTTGAT 113666
QY 1920 GTTCAAAATCTTGGCAATCGGCGGAAAACGG-CAATGCTTTTCAATTGAACATCATCTTCT 1978
Db 113665 GTTCAAAATCTTGGCAATCGGCGGAAAACGGCAATGCTTTCAATTGAACATCATCTTCT 113606
QY 1979 CTGTGAGGACTGCAAAAGGCTGCAAGGCGTGCACCTTAAACGTGAGAGCAACGCTTGA 2038
Db 113605 CTGTGAGGACTGCAAAAGGCTGCAAGGCGTGCACCTTAAACGTGAGAGCAACGCTTGA 113546
QY 2039 TTCTTCAACACCATGCAAGAGGTCCCAACCCAGCGCTCACCCAGGTTTGGGCACTTGG 2098
Db 113545 TTCTTCAACACCATGCAAGAGGTCCCAACCCAGCGCTCACCCAGGTTTGGGCACTTGG 113486
QY 2099 TCCCGCGCGCAAAACATCTGTGTGGTGGCACTGTGTGAAGAAAAGCCGACGCTCCG 2158
Db 113485 TCCCGCGCGCAAAACATCTGTGTGGTGGCACTGTGTGAAGAAAAGCCGACGCTCCG 113426
QY 2159 GGAACGTGGAAGGCGCCAGTGACTGCTTCTTGGCCAGGTTCCATCTGTGATGAGCAAC 2218
Db 113425 GGAACGTGGAAGGCGCCAGTGACTGCTTCTTGGCCAGGTTCCATCTGTGATGAGCAAC 113366
QY 2219 ATGCCACCATCATCTGTGATGAGAGCAGATATCCAGCTGGAAGAACGCTGATCATAC 2278
Db 113365 ATGCCACCATCATCTGTGATGAGAGCAGATATCCAGCTGGAAGAACGCTGATCATAC 113306
QY 2279 CGTCTCATGAGACATTTAAAGCTGGCTGAGAAAACAAAAGAAAGTACTGTGTGGGCTTA 2338
Db 113305 CGTCTCATGAGACATTTAAAGCTGGCTGAGAAAACAAAAGAAAGTACTGTGTGGGCTTA 113246
QY 2339 TGCAACAGAACTTTCCAGTTTGGCGCCCTGCTACATGTGACTCTCCGCGAGGCGAGGC 2398
Db 113245 TGCAACAGAACTTTCCAGTTTGGCGCCCTGCTACATGTGACTCTCCGCGAGGCGAGGC 113186
QY 2399 TCAATGATCCCAAGGAATGTAAGTGAATGATACCTTCCACGCTTACTACAGACG 2458
Db 113185 TCAATGATCCCAAGGAATGTAAGTGAATGATACCTTCCACGCTTACTACAGACG 113126
QY 2459 ATCCAGGTTTCCCTTGGCAACCAAGCGGACCGGCTGGGCTACACACACAGCGCTTGA 2518
Db 113125 ATCCAGGTTTCCCTTGGCAACCAAGCGGACCGGCTGGGCTACACACACAGCGCTTGA 113066
QY 2519 CCGGACCGGAGGATGAGTGAAGGACGCTGCGGACGCTCTTTTACCAGGATGATCCT 2578
Db 113065 CCGGACCGGAGGATGAGTGAAGGACGCTGCGGACGCTCTTTTACCAGGATGATCCT 113006
QY 2579 ATGACCTGATGATGATGATTCCTGGGTGAAGCCGTAATTTACTGACGCAACTTAAACTTT 2638
Db 113005 ATGACCTGATGATGATGATTCCTGGGTGAAGCCGTAATTTACTGACGCAACTTAAACTTT 112946
QY 2639 TCTTACACCGGCACTTAAATAATTTGACGGAAGCGCGGCGCACCCAAAACCTTGTGAG 2698
Db 112945 TCTTACACCGGCACTTAAATAATTTGACGGAAGCGCGGCGCACCCAAAACCTTGTGAG 112886

QY 2699 TCGAGGACCCCAACTGGGCTGATGGGCGGCACTTCAATCGCGTTGCTTAAATAATCCGCTTA 2758
Db 112885 TCGAGGACCCCAACTGGGCTGATGGGCGGCACTTCAATCGCGTTGCTTAAATAATCCGCTTA 112826
QY 2759 TCGAGGACCCGCGAGCGGTTTCAACCCCAATTACCGGATCCCATGATAGCCCTGATG 2818
Db 112825 TCGAGGACCCGCGAGCGGTTTCAACCCCAATTACCGGATCCCATGATAGCCCTGATG 112766
QY 2819 GTGATGTTGGAACATGCTTCTTGGGCGCCAAACGGAACCTCAACCGGTGACGCGTTC 2878
Db 112765 GTGATGTTGGAACATGCTTCTTGGGCGCCAAACGGAACCTCAACCGGTGACGCGTTC 112706
QY 2879 TATACCGCTGACAGATCTTGAATACTGGGAATCTCCGGAATATCACTTTGACCTCA 2938
Db 112705 TATACCGCTGACAGATCTTGAATACTGGGAATCTCCGGAATATCACTTTGACCTCA 112646
QY 2939 GTGATGCAAACTGTGTTCTGCTCTGATCTGTTCCGATAGGCTACATGTGGGAATGCC 2998
Db 112645 GTGATGCAAACTGTGTTCTGCTCTGATCTGTTCCGATAGGCTACATGTGGGAATGCC 112586
QY 2999 CCAACCTTTTTCACGCTTGGCATGAGAAATGGCGAAAGATCTGACGCTGATTTTCT 3058
Db 112585 CCAACCTTTTTCACGCTTGGCATGAGAAATGGCGAAAGATCTGACGCTGATTTTCT 112526
QY 3059 GTCCCAAGAGATTGGAACCGAATCCACAGATGAGTTTACTCATACGCAAGCTCTGACAGT 3118
Db 112525 GTCCCAAGAGATTGGAACCGAATCCACAGATGAGTTTACTCATACGCAAGCTCTGACAGT 112466
QY 3119 GCGATATGTGCTGACAAAGCTTGAAGGAAGCACTTCCGGCTTGGAGAGATTCAAGC 3178
Db 112465 GCGATATGTGCTGACAAAGCTTGAAGGAAGCACTTCCGGCTTGGAGAGATTCAAGC 112406
QY 3179 AGCTGGAATTTGGGCTACATGATTTACGCAACCGGAGTTGAGATTAACGCTCTGATGCT 3238
Db 112405 AGCTGGAATTTGGGCTACATGATTTACGCAACCGGAGTTGAGATTAACGCTCTGATGCT 112346
QY 3239 GGCCTGAGGCTGATGAGGCGCTGCGCGCGAGATGATCAACCAAGTTGCAAGAGAG 3298
Db 112345 GGCCTGAGGCTGATGAGGCGCTGCGCGCGAGATGATCAACCAAGTTGCAAGAGAG 112286
QY 3299 GATGGGTGACCTGCTGCTGCTGCGCGCGAGCTTCAATTGGCAACCAACGCGATCTAC 3358
Db 112285 GATGGGTGACCTGCTGCTGCTGCGCGCGAGCTTCAATTGGCAACCAACGCGATCTAC 112226
QY 3359 AAGAGCTCTCTTCCCGAGGCGGAGTCCGGGATTAACAGTCTGTATTAGCTCTGAC 3418
Db 112225 AAGAGCTCTCTTCCCGAGGCGGAGTCAAGGATTAACAGTCTGTATTAGCTCTGAC 112166
QY 3419 CTGTCCGATGAGATCCGAGGCAATTTTCCCTGAGTGGGATGCTGCTCGTTGCTG 3478
Db 112165 CTGTCCGATGAGATCCGAGGCAATTTTCCCTGAGTGGGATGCTGCTCGTTGCTG 112106
QY 3479 TGGATCGTATGATGATGCTGCGGTAAGTGAATTAATTAATTAATTAATTAATTAAT 3538
Db 112105 TGGATCGTATGATGATGCTGCGGTAAGTGAATTAATTAATTAATTAATTAATTAAT 112046
QY 3539 ACGATTAATACGCGATGAGTAACTGACAGTGAATGACAGGTTTCAATCCCTTTCCG 3598
Db 112045 ACGATTAATACGCGATGAGTAACTGACAGTGAATGACAGGTTTCAATCCCTTTCCG 111987
QY 3599 GCTTCAAGAGTGAACATTAATGAGATTAAGTCAATTAATTAATTAATTAATTAATTAAT 3658
Db 111986 GCTTCAAGAGTGAACATTAATGAGATTAAGTCAATTAATTAATTAATTAATTAATTAAT 111927
QY 3659 GTACAAATATCTTCCCAAAATCTTGAATGAGCAAAATTAACAGTTTATGTTTGA 3718
Db 111926 GTACAAATATCTTCCCAAAATCTTGAATGAGCAAAATTAACAGTTTATGTTTGA 111867
QY 3719 GCTTTGAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3778
Db 111866 GCTTTGAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 111807
QY 3779 ATGAGACATTAAGGACCTGCGCAACGATCTGCGGACATTTGGCGGAGCAACATTT 3838

Db 111806 ATGACCAATPAGGACCTCGGCGACGCAATCTGCGGACATTTGGCCGCGAAGCAACATT 111747
Qy 3839 GTGCGCGCGCACACTGTGCAACGCGTTTACGCTCTGTGCTCAAGAACACCAAGATGTG 3898
Db 111746 GTGCGCGCGCACACTGTGCAACGCGTTTACGCTCTGTGCTCAAGAACACCAAGATGTG 111687
Qy 3899 GATGCGCAAGTGTGATGATGATCCAGATCTGAAAGGCACTTTGAAACTGCGCGCATG 3958
Db 111686 GATGCGCAAGTGTGATGATGATCCAGATCTGAAAGGCACTTTGAAACGGGTGGTATG 111627
Qy 3959 TTCCAGATCATCGTCCGCGCAGCGCATGTGATCATGTTTCAAGAACTCGATGACGA 4018
Db 111626 TTCCAGATCATCGTCCGCGCAGCGCATGTGATCATGTTTCAAGAACTCGATGACGA 111567
Qy 4019 ACCTCCAAAGACATCGCTGTGTCCACAGACAGCTCAAAAGATTGTGTGCTAAACAGCC 4078
Db 111566 ACCTCCAAAGACATCGCTGTGTCCACAGACAGCTCAAAAGATTGTGTGCTAAACAGCC 111507
Qy 4079 AACTGTTGAGCGGTGCTGTAAGATTTGGCGGACATTTTGGTCCGCTGATTCGAATC 4138
Db 111506 AACTGTTGAGCGGTGCTGTAAGATTTGGCGGACATTTTGGTCCGCTGATTCGAATC 111447
Qy 4139 TTGTTGTGTGCGGTCTGCTCATGTGCTCAACATGTGTGTTGCGCAGAGATCTGTTC 4198
Db 111446 TTGTTGTGTGCGGTCTGCTCATGTGCTCAACATGTGTGTTGCGCAGAGATCTGTTC 111387
Qy 4199 GGTGCGCAATCACTGTGTGAGATGTTCCCTCAGATCAGCGGTGTTGCTGATGATCAAC 4258
Db 111386 GGTGCGCAATCACTGTGTGAGATGTTCCCTCAGATCAGCGGTGTTGCTGATGATCAAC 111327
Qy 4259 CTGATGAGCATCTGCGCGCTTGGGTCTTGGCCAGTGTGTGTTGTTTACCGGACCAAG 4318
Db 111326 CTGATGAGCATCTGCGCGCTTGGGTCTTGGCCAGTGTGTGTTGTTTACCGGACCAAG 111267
Qy 4319 CGTTTGTGTGCAATGATGTTCTCGGCGCGCGCATTTGATGATGCGATGTTCCCAAC 4378
Db 111266 CGTTTGTGTGCAATGATGTTCTCGGCGCGCGCATTTGATGATGCGATGTTCCCAAC 111207
Qy 4379 CTGTTTACGCGCTACGACGTGCGCGCACATGACCGCGCGCAAAATGCAATGTGTTC 4438
Db 111206 TTGGTGAAACGCGCTACGACGTGCGCGCACATGACCGCGCGCAAAATGCAATGTGTTC 111147
Qy 4439 CTGTTTGTGTGATGATGTTCTCAAGCTGTGTACCAAGGCAACGTTCTCTGTGCTGTG 4498
Db 111146 CTGTTTGTGTGATGATGTTCTCAAGCTGTGTGTACCAAGGCAACGTTCTCTGTGCTGTG 111087
Qy 4499 GTCTCTTGTGATCTGTGCAACGATCGAAGATTCTGCAACAGGCACTCAAGGCACTGCA 4558
Db 111086 GTCTCTTGTGATCTGTGCAACGATCGAAGATTCTGCAACAGGCACTCAAGGCACTGCA 111027
Qy 4559 GACTTCCTGATCAACCCAGTGTGATCTGTGCTGCTGACCGGCTTCTTACGTTCAATGTCT 4618
Db 111026 GACTTCCTGATCAACCCAGTGTGATCTGTGCTGCTGACCGGATTCCTTACATTCATCGCG 110967
Qy 4619 ATTGTGTCAGCAATGCGCTGGGTGGGTGACTGTCTGTGCAACGAGTGTGCAAGGACCTTAT 4678
Db 110966 ATTGTGTCAGCAATGCGCTGGGTGGGTGACTGTCTGTGCAACGAGTGTGCAAGGACCTTAT 110907
Qy 4679 GATTTGCGTGTGCAAGTGTGCGGCTGTGTTTCTGCTGTGCTGATCAACCAATCGTTATC 4738
Db 110906 GATTTGCGTGTGCAAGTGTGCGGCTGTGTTTCTGCTGTGCTGATCAACCAATCGTTATC 110847
Qy 4739 ACTGTGTCAGCAAGTCTTCCGCGCAATGATGATGATGATGTTCAACCAAGGATGATCC 4798
Db 110846 ACTGTGTCAGCAAGTCTTCCGCGCAATGATGATGATGATGTTCAACCAAGGATGATCC 110787
Qy 4799 TTGATCTTGTGCAACCGCATTCATGCGCAATATGTGCGCAGGCTGACAGATGTTTGGCAGTG 4858
Db 110786 TTGATCTTGTGCAACCGCATTCATGCGCAATATGTGCGCAGGCTGACAGATGTTTGGCAGTG 110727
Qy 4859 TTCTTCTGAGCAAGATGAAAAAGCTCAAGGCGCTTGAGGCTTCAAGTGTCTCGGCT 4918

Db 110726 TTCTTCTGAGCAAGATGAAAAAGCTCAAGGCGCTTGAGGCTTCAAGTGTCTCGGCT 110667
Qy 4919 GTTCTTGTGATTAACAGAGCTTCCGATCTTCCGTTGTGAACCTTCCGCTGCGCTGCGCTTC 4978
Db 110666 GTTCTTGTGATTAACAGAGCTTCCGATCTTCCGTTGTGAACCTTCCGCTGCGCTGCGCTTC 110607
Qy 4979 TACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5038
Db 110606 TACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 110547
Qy 5039 GCAGTTGCTGTTGGCGCTGACAGTGTCTTGGGTGTGTTTCTTATGATGCTCCAGATATG 5098
Db 110546 GCAGTTGCTGTTGGCGCTGACAGTGTCTTGGGTGTGTTTCTTATGATGCTCCAGATATG 110487
Qy 5099 GTGATGTTTGTGTTTGGCGGTGATGATCTTGTGATGATGATGATGATGATGATGATGATGAT 5158
Db 110486 GTGATGTTTGTGTTTGGCGGTGATGATCTTGTGATGATGATGATGATGATGATGATGATGAT 110427
Qy 5159 TATGGCTTTTACTTGTGCGCGCAACGCGACATGATTCAGATGCAACGCGTGTCCA 5218
Db 110426 TATGGCTTTTACTTGTGCGCGCAACGCGACATGATTCAGATGCAACGCGTGTCCA 110367
Qy 5219 GTGCTGACAGAACGACCAAGCCGAAAGCAGAACCCGACAGATTTTCAACGATTC 5278
Db 110366 GTGCTGACAGAACGACCAAGCCGAAAGCAGAACCCGACAGATTTTCAACGATTC 110307
Qy 5279 ACATCATCAGGACCTTTGACCGGTGAGCTATGCACTGACAGCGTCAAGCATGCC 5338
Db 110306 ACATCATCAGGACCTTTGACCGGTGAGCTATGCACTGACAGCGTCAAGCATGCC 110247
Qy 5339 ATGTTTGCAGGGAAGGCTTGGCTCAGTGTGAGTGTGCTGCTCCCAACCAAGGCGGACGCG 5398
Db 110246 ATGTTTGCAGGGAAGGCTTGGCTCAGTGTGAGTGTGCTGCTCCCAACCAAGGCGGAGTGA 110187
Qy 5399 GTTTCACAGTAGGCGGAAGATCGTGTGCTTCCCATCTGTGCTGACGCTTTCGACATC 5458
Db 110186 GTTTCACAGTAGGCGGAAGATCGTGTGCTTCCCATCTGTGCTGACGCTTTCGACATC 110127
Qy 5459 GCGACTAAGCTGAGATGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5518
Db 110126 GCGACTAAGCTGAGATGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 110067
Qy 5519 GTTAACTCAAGGACCGACCTTTAACCCGCTGAAGAAAGCGGAGTGAAGTCAAAAGA 5578
Db 110066 GTTAACTCAAGGACCGACCTTTAACCCGCTGAAGAAAGCGGAGTGAAGTCAAAAGA 110007
Qy 5579 GGGAGAGCTGCTGTGAAATTCGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5638
Db 110006 GGGAGAGCTGCTGTGAAATTCGATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 109947
Qy 5639 ACGCGATTTGTTTGAATTAACAAGAAACCGGACCTGTAAACATTAACGTTTGGGCG 5698
Db 109946 ACGCGATTTGTTTGAATTAACAAGAAACCGGACCTGTAAACATTAACGTTTGGGCG 109887
Qy 5699 GAAATTTGAAGCGGAGCCCACTGTCTCAACGTCGAAAGAAAGAAAGCGGTGCAACA 5758
Db 109886 GAAATTTGAAGCGGAGCCCACTGTCTCAACGTCGAAAGAAAGAAAGCGGTGCAACA 109827
Qy 5759 CCATAGTTTGAACCTTGAATGTTGCAACAGATTTAGCTAGGAGAGCTGATGATGATGATGATGAT 5818
Db 109826 CCATAGTTTGAACCTTGAATGTTGCAACAGATTTAGCTAGGAGAGCTGATGATGATGATGATGAT 109767
Qy 5819 ATCTTTGACACCGGATACCGGTGAGATTTGAACTTTTAACTTTTCAACAGATGATGCC 5878
Db 109766 ATCTTTGACACCGGATACCGGTGAGATTTTAACTTTTAACTTTTCAACAGATGATGCC 109707
Qy 5879 TCGGTGTACCTGTGTGTGTGCAACCCCGCAATCTTCAACCCCATGAGACATGTTGATGATGATGAT 5938
Db 109706 TCGGTGTACCTGTGTGTGTGCAACCCCGCAATCTTCAACCCCATGAGACATGTTGATGATGATGAT 109647
Qy 5939 GCAGTAGGCTTGAATTTTGGCGGCTG 5967
Db 109646 GCAGTAGGCTTGAATTTTGGCGGCTG 109618

Db 961 ATACACATACAGCGCAATTGCTTGCTAAAGAGAAAATGAGCGTCAAGCGCTAACCG 1020
Qy 1021 ATTTGTTCAATGCGATGCTCTCGCTGATCATAGGGCTCCGGACGCTGGCGCTTTGC 1080
Db 1022 ATTTGTTCAATGCGATGCTCTCGCTGATCATAGGGCTCCGGACGCTGGCGCTTTGC 1080
Qy 1081 TTGCTGCGGACGCTGCGGGGAGCGCATGTTGATGATGCGCGAGCGGCTGCTATTGG 1140
Db 1081 TTGCTGCGGACGCTGCGGGGAGCGCATGTTGATGATGCGCGAGCGGCTGCTATTGG 1140
Qy 1141 CGGATGGAACGGTGCATCTAGCTCGTTCGAACAACGCTTTTTCATCAGCGAGCCATGG 1200
Db 1141 CGGATGGAACGGTGCATCTAGCTCGTTCGAACAACGCTTTTTCATCAGCGAGCCATGG 1200
Qy 1201 AAGCGCCCGGAATGCCAGACGGTGAATCAATTTTGCGCTTTTGAACGTCAACGTCACCG 1260
Db 1201 AAGCGCCCGGAATGCCAGACGGTGAATCAATTTTGCGCTTTTGAACGTCAACGTCACCG 1260
Qy 1261 ATGAGATGCGCGCTGCGCGGATGGGGGCGCATCCCGGGGCGACAGACATAGCGA 1320
Db 1261 ATGAGATGCGCGCTGCGCGGATGGGGGCGCATCCCGGGGCGACAGACATAGCGA 1320
Qy 1321 GTCAAGTTCGTGACACACGTCGCGAGGGGTATGACGCTTATCGACGCGACCTTCACACT 1380
Db 1321 GTCAAGTTCGTGACACACGTCGCGAGGGGTATGACGCTTATCGACGCGACCTTCACACT 1380
Qy 1381 CAACCGTCCGCTTAAATTTCTCGTCTTGGCGATCACGAATCGCTTAATCCAAACCTTG 1440
Db 1381 CAACCGTCCGCTTAAATTTCTCGTCTTGGCGATCACGAATCGCTTAATCCAAACCTTG 1440
Qy 1441 CAATTTTGCTGCTTGACTCAACGCGGACGGTGAAGGCTCAATTTAGGCTATCAAG 1500
Db 1441 CAATTTTGCTGCTTGACTCAACGCGGACGGTGAAGGCTCAATTTAGGCTATCAAG 1500
Qy 1501 TACTTTAGTACGAGTAAATCTATCCTGATTTTAAAGAGTCCCAACATGAAATCACTA 1560
Db 1501 TACTTTAGTACGAGTAAATCTATCCTGATTTTAAAGAGTCCCAACATGAAATCACTA 1560
Qy 1561 TCTGCAAGACGAGCAAGAAAGTGGCAAAAGCAAGTTCAGTCTTAATCGACCTTGGCA 1620
Db 1561 TCTGCAAGACGAGCAAGAAAGTGGCAAAAGCAAGTTCAGTCTTAATCGACCTTGGCA 1620
Qy 1621 ACAAGGGTGAACCTTGCGGCTTGGCAACGAGATCCTCACAGTATCTACCAAGAGC 1680
Db 1621 ACAAGGGTGAACCTTGCGGCTTGGCAACGAGATCCTCACAGTATCTACCAAGAGC 1680
Qy 1681 TCAATTCGATGATGAGAGTGGGGAAGTGTCAATTAAGATGCGAAGCATTTCTTGG 1740
Db 1681 TCAATTCGATGATGAGAGTGGGGAAGTGTCAATTAAGATGCGAAGCATTTCTTGG 1740
Qy 1741 ATGAATAGTGGGACTTAACCCGTGACGATGAAAAAGCTTCTTTAAACCATTTGCAAG 1800
Db 1741 ATGAATAGTGGGACTTAACCCGTGACGATGAAAAAGCTTCTTTAAACCATTTGCAAG 1800
Qy 1801 AGTTTACATGACACATCGATCGATGAGAGAGGCTTCAACGCCAGATGGTGAAC 1860
Db 1801 AGTTTACATGACACATCGATCGATGAGAGAGGCTTCAACGCCAGATGGTGAAC 1860
Qy 1861 CTGATTCATACGAGAGCTGACAGATGAGGCAAGATCGCTGACAGATCGTTGAAG 1920
Db 1861 CTGATTCATACGAGAGCTGACAGATGAGGCAAGATCGCTGACAGATCGTTGAAG 1920
Qy 1921 TTCAAAATCTTGGCATCGCGGAAACGGCAATCGCTTTCAATTGAACATCTTCTCT 1980
Db 1921 TTCAAAATCTTGGCATCGCGGAAACGGCAATCGCTTTCAATTGAACATCTTCTCT 1980
Qy 1981 GTCAAGACTGAACAAAGGCTCAGGCGCTGACCTTAACATGAGAGAGCAAGCGCTGAT 2040
Db 1981 GTCAAGACTGAACAAAGGCTCAGGCGCTGACCTTAACATGAGAGAGCAAGCGCTGAT 2040
Qy 2041 GTTCAACACCATGAGAGGTCCCAACCGCGCTCACCGAGGTTTGGGCACTTTGTC 2100
Db 2041 GTTCAACACCATGAGAGGTCCCAACCGCGCTCACCGAGGTTTGGGCACTTTGTC 2100

Qy 2101 CCGCGCGCAAAACATCGTCTTGGTGCACTGTGAAGAAAAGCCGACCGCATCCGCGG 2160
Db 2101 CCGCGCGCAAAACATCGTCTTGGTGCACTGTGAAGAAAAGCCGACCGCATCCGCGG 2160
Qy 2161 AACTGTGAAGGCGCAGTACGCTTCTTGGCCAGGTTTCATCTGTAGATCACAAAT 2220
Db 2161 AACTGTGAAGGCGCAGTACGCTTCTTGGCCAGGTTTCATCTGTAGATCACAAAT 2220
Qy 2221 GCCACCATCATGTTGATGAGAGCAGATATCAAGCTGAAAAAGCTGATCACTACCG 2280
Db 2221 GCCACCATCATGTTGATGAGAGCAGATATCAAGCTGAAAAAGCTGATCACTACCG 2280
Qy 2281 TCTCATGAGCAATTAAGCTCGCTGAGAAACAAAAGGAAAGTCTGTGGGCTATG 2340
Db 2281 TCTCATGAGCAATTAAGCTCGCTGAGAAACAAAAGGAAAGTCTGTGGGCTATG 2340
Qy 2341 CACACGAACCTTCCAGTTTGGCGCTGCTGCAATGAGACTCTCCGACGGGCAAGCTC 2400
Db 2341 CACACGAACCTTCCAGTTTGGCGCTGCTGCAATGAGACTCTCCGACGGGCAAGCTC 2400
Qy 2401 AATGATCCCAACGGAATGATCGATGAGATACCTTCACGCTACTACAGACGAT 2460
Db 2401 AATGATCCCAACGGAATGATCGATGAGATACCTTCACGCTACTACAGACGAT 2460
Qy 2461 CCAAGTTTCCCTTGGCACACAAAGCGCACCGGCTGGGCTCACACCAACCGCTTGAC 2520
Db 2461 CCAAGTTTCCCTTGGCACACAAAGCGCACCGGCTGGGCTCACACCAACCGCTTGAC 2520
Qy 2521 GGAACGACGAGATTCAGATGAGACGACCTGCGCGAGCGCTCTTACCGGATGATCTAT 2580
Db 2521 GGAACGACGAGATTCAGATGAGACGACCTGCGCGAGCGCTCTTACCGGATGATCTAT 2580
Qy 2581 GACCTGATGATGATGATTCGCTGAGAGCGGATTTACTAGAGGACACTTAACTTTTC 2640
Db 2581 GACCTGATGATGATGATTCGCTGAGAGCGGATTTACTAGAGGACACTTAACTTTTC 2640
Qy 2641 TACACCGGCAACCTTAAATTAACGAAAAGCGCGCGCACCCAAACCTTGTGAAAGTC 2700
Db 2641 TACACCGGCAACCTTAAATTAACGAAAAGCGCGCGCACCCAAACCTTGTGAAAGTC 2700
Qy 2701 GAGAACCCAACTGGGCTGATGGGCGGCAATTCATGCGCGTGGCTTAAATTCGCTTATC 2760
Db 2701 GAGAACCCAACTGGGCTGATGGGCGGCAATTCATGCGCGTGGCTTAAATTCGCTTATC 2760
Qy 2761 GACGAGCCCGCAGCGGTTTCAACCCCATTAACCGGATCCCATGATGACGCTGATGT 2820
Db 2761 GACGAGCCCGCAGCGGTTTCAACCCCATTAACCGGATCCCATGATGACGCTGATGT 2820
Qy 2821 GATGTTGGAACATGCTTCTTGGGCGCAACCGGAAAACCTTACCGGTGACGCTTCTA 2880
Db 2821 GATGTTGGAACATGCTTCTTGGGCGCAACCGGAAAACCTTACCGGTGACGCTTCTA 2880
Qy 2881 TACCGCTGACGATCTTGAAGAACTGCGGGAATCTCCGGAATCAACCTTGAACCTCAGT 2940
Db 2881 TACCGCTGACGATCTTGAAGAACTGCGGGAATCTCCGGAATCAACCTTGAACCTCAGT 2940
Qy 2941 GATGCAAAACCTGTTCTGCTCTGATCTCGTTCCGAGTGGCTACATGATGGAATGCCCC 3000
Db 2941 GATGCAAAACCTGTTCTGCTCTGATCTCGTTCCGAGTGGCTACATGATGGAATGCCCC 3000
Qy 3001 AACCTTTTACGCTTCCGATGAGAAACCTGCGGAAAGTCTGACGTGCTGATTTTCTGT 3060
Db 3001 AACCTTTTACGCTTCCGATGAGAAACCTGCGGAAAGTCTGACGTGCTGATTTTCTGT 3060
Qy 3061 CCAAGAGATTGAGACGATCCAGATGAGGTTACTACTAGCAAGCTTGAACAGTGC 3120
Db 3061 CCAAGAGATTGAGACGATCCAGATGAGGTTACTACTAGCAAGCTTGAACAGTGC 3120
Qy 3121 GGATATGTCGTGACAAACCTTGAAGAAAGCACTTCCGCGTTCGAGAGATTGAGAG 3180
Db 3121 GGATATGTCGTGACAAACCTTGAAGAAAGCACTTCCGCGTTCGAGAGATTGAGAG 3180

QY 3181 CTGATTTTCGGCATGAATTCTAGCCACCGAGTTGAGTAACGGTTCTGATCCCTGG 3240
 DB 3181 CTGATTTTCGGCATGAATTCTAGCCACCGAGTTGAGTAACGGTTCTGATCCCTGG 3240
 QY 3241 CTCCTGGGCTGATGAGGGGCTGCCCGGAGATGATCAACCAAGTTGACAGGAAGA 3300
 DB 3241 CTCCTGGGCTGATGAGGGGCTGCCCGGAGATGATCAACCAAGTTGACAGGAAGA 3300
 QY 3301 TGGGTGACACTGCTGCTGCTCCCGCAAGCTTCAATTGGCGAACACGCGATCTACCA 3360
 DB 3301 TGGGTGACACTGCTGCTGCTCCCGCAAGCTTCAATTGGCGAACACGCGATCTACCA 3360
 QY 3361 GAGCTCTTCTCCAGAGGGGAGTCCGGGGTAACTCAAGTCTGTTAGTTTAACTT 3420
 DB 3361 GAGCTCTTCTCCAGAGGGGAGTCCGGGGTAACTCAAGTCTGTTAGTTTAACTT 3420
 QY 3421 GTCCGAGTGAATCCGAGGCAATATTTCCCTGAGTGGAGTGTCCGTTGTCTGTG 3480
 DB 3421 GTCCGAGTGAATCCGAGGCAATATTTCCCTGAGTGGAGTGTCCGTTGTCTGTG 3480
 QY 3481 GATCGTATGATGATCGTCCGCTAGCTAGGTAAGTAAACCTGCGCAATTAGTATCGCGAC 3540
 DB 3481 GATCGTATGATGATCGTCCGCTAGCTAGGTAAGTAAACCTGCGCAATTAGTATCGCGAC 3540
 QY 3541 GATTAATACGCAATGATGATTAATCTGAGGTATGACAGGTTTCAATTCCTTTCCGGGC 3600
 DB 3541 GATTAATACGCAATGATGATTAATCTGAGGTATGACAGGTTTCAATTCCTTTCCGGGC 3600
 QY 3601 CTTCAAGGTGACACTATTGAGAGATTAAGTCAATTAATAAGGCTTTTGGCGCAATTGT 3660
 DB 3601 CTTCAAGGTGACACTATTGAGAGATTAAGTCAATTAATAAGGCTTTTGGCGCAATTGT 3660
 QY 3661 ACAAAATACTTCCGCAAAATCCCTTGAATC 3687
 DB 3661 ACAAAATACTTCCGCAAAATCCCTTGAATC 3687
 RESULT 4
 ID AAH67869 standard; DNA; 1983 BP.
 AC AAH67869;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE C glutamicum coding sequence fragment SEQ ID NO: 2904.
 XX
 KM Corynebacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; ds.
 XX
 OS Corynebacterium glutamicum.
 OS
 PN EP1108790-A2.
 XX
 PD 20-JUN-2001.
 XX
 PF 18-DEC-2000; 2000BP-00127688.
 XX
 PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H,
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 XX
 DR WPI: 2001-376931/40.
 DR P-PSDB; AAG92650.
 XX
 PT Novel polynucleotides derived from Corynebacterium, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.

XX
 PS Claim 8, SEQ ID NO 2904; 246pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the Corynebacterium glutamicum. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of corynebacterium, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC Corynebacterium, and identifying a homologue of a gene derived from
 CC corynebacterium. Corynebacterium are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 SQ Sequence 1983 BP; 410 A; 508 C; 558 G; 507 T; 0 U; 0 Other;
 Query Match 31.4%; Score 1874.2; DB 5; Length 1983;
 Best Local Similarity 96.6%; Pred. No. 0;
 Matches 1915; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
 QY 3779 ATGAGCATAAAGACCTCGCGCAACGCATCTCGCGCAATTTGGACCAATTT 3838
 DB 1 ATGAGCATAAAGACCTCGCGCAACGCATCTCGCGCAATTTGGACCAATTT 60
 QY 3839 GTCCGCGCGCACTGTGCAACGCCCTTACGCTCGTCAAAAGACCAAGGATGTG 3898
 DB 61 GTCCGCGCGCACTGTGCAACGCCCTTACGCTCGTCAAAAGACCAAGGATGTG 120
 QY 3899 GATCGCAAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3958
 DB 121 GATCGCAAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
 QY 3959 TTCAAGATCATCTCGCGCGCAACGCATCTCGCGCAATTTGGACCAATTT 4018
 DB 181 TTCAAGATCATCTCGCGCGCAACGCATCTCGCGCAATTTGGACCAATTT 240
 QY 4019 AACTCCAAAGACATCGCTGTGTCCAGAGAGCGCTCAAAAGTGTGGCTTAACAAGCC 4078
 DB 241 AACTCCAAAGACATCGCTGTGTCCAGAGAGCGCTCAAAAGTGTGGCTTAACAAGCC 300
 QY 4079 AACTGTTCACGCGGTCTGTGAAGTATTGGCGCAATTTTGTCCGCTGATTCATTC 4138
 DB 301 AACTGTTCACGCGGTCTGTGAAGTATTGGCGCAATTTTGTCCGCTGATTCATTC 360
 QY 4139 TTGGTGTGCGGTCTGTCTATGCTTAACAAGTGTGTGTGTGTGTGTGTGTGTGTGT 4198
 DB 361 TTGGTGTGCGGTCTGTCTATGCTTAACAAGTGTGTGTGTGTGTGTGTGTGTGTGT 420
 QY 4199 GGTCCGAATCACTGTGTGAAGTGTCTCCAGATCAAGCGGTGTGTGTGTGTGTGTGTGT 4258
 DB 421 GGTCCGAATCACTGTGTGAAGTGTCTCCAGATCAAGCGGTGTGTGTGTGTGTGTGTGT 480
 QY 4259 CTGATGATCTGCGCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4318
 DB 481 CTGATGATCTGCGCGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
 QY 4319 CGTTTCGT 4378
 DB 541 CGTTTCGT 600
 QY 4379 CTGTTTAAAGGCTAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4438
 DB 601 TTGGTGAACGCTAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
 QY 4439 CTGTTTGT 4498
 DB 661 CTGTTTGT 720
 QY 4499 GTCTCTGT 4558
 DB 721 GTCTCTGT 780


```

QY 4559 GACTTCCTGATGACCCAGGTTGACTGCTGCTCAGCGGCTTCTTACGTTCAATTGCT 4618
DB 781 GACTTCCTGATCACTCAGTGTGCTGAGCTGTGCTGCTACCGGANTCTTACATTCATCGCC 840
QY 4619 ATTGATCAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4678
DB 841 ATTGATCAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 4679 GATTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4738
DB 901 GATTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 4739 ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4798
DB 961 ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 4799 TTTCATCTTGCAGACCGCATCTCATGCTCAATATCGCGAGGCTGAGCATGTTTGGCAGTG 4858
DB 1021 TTTCATCTTGCAGACCGCATCTCATGCTCAATATCGCGAGGCTGAGCATGTTTGGCAGTG 1080
QY 4859 TTCTTCTTACGAGAGGTTGAAAAGCTCAAGGCTTTCAGGCTTTCAGGCTTTCAGGCT 4918
DB 1081 TTCTTCTTACGAGAGGTTGAAAAGCTCAAGGCTTTCAGGCTTTCAGGCTTTCAGGCT 1140
QY 4919 GTTCTTGTGATTAACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4978
DB 1141 GTTCTTGTGATTAACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 4979 TACATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5038
DB 1201 TACATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 5039 GCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5098
DB 1261 GCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 5099 GTTCATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5158
DB 1321 GTTCATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
QY 5159 TATGAGCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5218
DB 1381 TATGAGCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
QY 5219 GTGCTGCTGAGAAAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5278
DB 1441 GTGCTGCTGAGAAAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
QY 5279 ACCATCATCAGGACCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5338
DB 1501 ACCATCATCAGGACCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
QY 5339 ATGTTTGCAGGAGAAAGCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5398
DB 1561 ATGTTTGCAGGAGAAAGCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
QY 5399 GTTTCACAGTACGCGAGAAAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5458
DB 1621 GTTTCACAGTACGCGAGAAAGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680
QY 5459 CGCAGTACGCTGAGGATGCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5518
DB 1681 CGCAGTACGCTGAGGATGCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
QY 5519 GTTAACTCAAGGAGCGAGCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5578
DB 1741 GTTAACTCAAGGAGCGAGCTTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
QY 5579 GGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5638
DB 1801 GGGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860

```

```

QY 5639 AGCCGATGTTGTTGTTGCAATTACAGAAACCGAGCTGTAAACATTCAGGTTGGGC 5698
DB 1861 AGCCGATGTTGTTGTTGCAATTACAGAAACCGAGCTGTAAACATTCAGGTTGGGC 1920
QY 5699 GAAATTGAAGGCGGAGCCCACTGCTCAACGTCGCAAGAAAGAAAGCGGTGCGACCAACA 5758
DB 1921 GAAATTGAAGGCGGAGCCCACTGCTCAACGTCGCAAGAAAGAAAGCGGTGCGACCAACA 1980
QY 5759 CCA 5761
DB 1981 CCA 1983

RESULT 5
AAF87497
ID AAF87497 standard; DNA, 1656 BP.
XX
XX AAF87497;
AC
XX
XX 09-JUL-2001 (first entry)
DT
XX
DE Corynebacterium thermaminogenes scrB nucleotide sequence.
XX
KW Corynebacterium; thermophilic; amino acid biosynthesis; enzyme;
KW thermotolerant; aceA; accBC; dtsR; dtsR2; pfk; scrB; gluABCD; pdhA; pc;
KW ppc; acn; icd; lpd; odhA; ds.
XX
OS Corynebacterium thermaminogenes.
XX
FH Key Location/Qualifiers
FT CDS 309..1598
FT /tag= a
FT /product= "scrB protein"
XX
XX WO200125447-A1.
XX
XX 12-APR-2001.
PD
XX
XX 04-OCT-2000; 2000WO-JP006913.
PF
XX
XX 04-OCT-1999; 99JP-00282716.
PR 01-NOV-1999; 99JP-00311147.
PR 21-APR-2000; 2000JP-00120687.
XX
XX (AJIN ) AJINOMOTO CO INC.
XX
XX Hirano S, Nonaka G, Matsuzaki Y, Akiyoshi N, Nakamura K;
XX Kimura E, Osumi T, Matsui K, Kawahara Y, Kurashashi O, Nakamatsu T;
XX Sugimoto S;
XX WPI; 2001-300170/31.
XX P-PSDB; AAB83190.
XX
XX Proteins and their DNA useful for microbial production of L-amino acids.
XX
XX Claim 27, Page 199-201, 215pp; Japanese.
XX
XX The present sequence is provided in a specification relating to genes
XX encoding thermophilic amino acid biosynthesis system enzymes of the
XX thermotolerant bacterium Corynebacterium thermaminogenes. The novel
XX proteins retain at least 30% isocitrate lyase activity after heating at
XX 500C for 5 minutes. DNA fragments encoding the enzymes were isolated from
XX a Corynebacterium thermaminogenes chromosomal DNA plasmid library by
XX PCR. The DNA may be used for developing strains of amino acid producing
XX microorganisms
XX
SQ Sequence 1656 BP; 392 A; 456 C; 432 G; 376 T; 0 U; 0 Other;
XX
Query Match 24.4%; Score 1457; DB 5; Length 1656;
Best Local Similarity 93.2%; Pred. No. 0;
Matches 1545; Conservative 0; Mismatches 110; Indels 2; Gaps 2;
QY 2031 ACCTCGATTCTTCAACACCATCGAAGAGTCCCAACCCAGCGCTCACCCAGGGTTGG 2090

```

```
Db 1 AGGCCGATTTCTTCAACTATCGAAGAGAGTCCCAACCCAGCTTTGACCCAGGGCTTGG 60
Qy 2091 GCATTTTGTCCCGCGCGCAAAACATCGTGTGTGTGGCAACTGTGTGAAGAAAAGCCGAC- 2149
Db 61 GTACTTTGTCCCGCGCGCAAAATATCGTGTGTGTGGCAACTGTGTGCAAGAAAAGCAGACA 120
Qy 2150 GGCATTCGCGGAACTGTGTGAAGAGCCCACTGACTGCTTCTTGCCAGGTTTCCATCTGTAG 2209
Db 121 GGCATTCGCGGAACTGTGTGAAGAGTCCAGTACTGCTTCTTGCCAGGTTTCCATCTGTGA 180
Qy 2210 ATGCAACAATGCAACCATCATCGTTGTGAAGAGAGAGTATCCAGCTGTGAAGAAACGT 2269
Db 181 ATGCAACAAGAGCCACCATCATCGTTGTGAAGAGAGAGATCCAGCTGTGAAGAAATGCT 240
Qy 2270 GATCACTACCGTCTCATGTGAGCAATTAAGCTGCGCTAGAAAACAAAAGAAAGTACTGT 2329
Db 241 GACCAATTACGTCATGTGAGCAATTAAGCTGCGCTAGAAAACAAAAGAAAGTACTGT 300
Qy 2330 GTGGGGCTTATGACACAGAACTTTTCCAGTTTGGCGCTGCGTACCATGTGACTCTCCGC 2389
Db 301 GTGGGGCTTATGACACAGAACTTTTCCAGTTTGGCGCTGCGTACCATGTGACTCTCCGC 360
Qy 2390 AGGGAGGCTCAATATCCCAACGGAATGTACGTGATGAGATACCTCCACGCTACT 2449
Db 361 AGGGAGACTCAATATATCCCAATGGAATGTACGTGATGAGATACCTCCACGCTACT 420
Qy 2450 ACCAGACGATCCAGGTTTCCCTTTGACACAAAGCGCAACGGCTGGGCTCAACACACA 2509
Db 421 ACCAGACGATCCAGGTTTCCCTTTGACACAAAGCGCAACGGTTGGGCTCAACACACA 480
Qy 2510 CGCGCTTACCGGAGACGAGCGATGTGACATGCCACCTGCGCGAGCTCTTACCCGG 2569
Db 481 CGCGCTTACCGGAGACGAGCGATGTGACATGCCACCTGCGCGAGCTCTTACCCGG 540
Qy 2570 ATGCATCTATGACCTGTGATGATCTATTCGCTGTGAGCGCTATTTACTGACGCGACAC 2629
Db 541 ATGCATCTATGACCTGTGATGATCTATTCGCGGAGAGCCGTATTTCTGACGCGACAC 600
Qy 2630 TTTAACTTTTCTTACACCGGCAACTTAAATTTGACGGAAGGCGCGCCACCCAAAAC 2689
Db 601 TTTAACTTTTCTTACACCGGCAACTTAAATTTGACGGAAGGCGCGCCACCCAAAAC 660
Qy 2690 TTGTGAAGTGAAGACCCCACTGGGCTGATGTGGGAGATTCATGCGCTGCGCTTAA 2749
Db 661 TTGTGAAGTGAAGACCCCACTGGGCTGATGTGGGAGATTCATGCGCTGCGCTTAA 720
Qy 2750 ATCCGCTTATCGACGGAACCGCGACGCGGTTTTCACACCCCATTAACGCGATCCATGATCA 2809
Db 721 ATCCGCTTATCGACGGAACCGCGACGCGGTTTTCACACCCCATTAACGCGATCCATGATCA 780
Qy 2810 GCCCTGATGTGATGTGTGAAACATGTGTTCTTGGGAGCCCAACGCGAAAACCTCACCGGTG 2869
Db 781 GCCCTGATGTGATGTGTGAAAGATGTGTTCTTGGGAGCTCACGCGAAAACCTCACCGGTG 840
Qy 2870 CAGGGGTTCTATAACCGCTGACAGATCTTGAATCTGGGAAATCTCCGCTGAAATCACT 2929
Db 841 CAGGGGTTCTATAACCGCTGACAGATCTTGAATCTGGGAAATCTCCGCTGAAATCACT 900
Qy 2930 TTGACCTGATGATGACAAACCTGGTCTGCTCTGATCTCGTTCCGATGCTCATGAT 2989
Db 901 TTGACCTGATGATGACAAACCTGGTCTGCTCTGATCTCGTTCCGCTGCTCATGAT 960
Qy 2990 GGGAAATGCCCAACTTTTACGCTTCCGATGAAAGAAATGCGAGATCTCGACGTC 3049
Db 961 GGGAAATGCCCAACTTTTACGCTTCCGATGAAAGAAATGCGAGATCTCGACGTC 1020
Qy 3050 TGATTTCTGTCACAAAGATTTGAGCCGAATCCAGATGAGGTTACTCATGACGAAGT 3109
Db 1021 TGATTTCTGTCACAAAGATTTGAGCCGAATCCAGATGAGGTTACTCATGACGAAGT 1080
Qy 3110 CTGACAGTGCAGATATGTCTGTCACAAAGCTTGAAGAAAGCACTTCCGCTCTTGCAG 3169
```

```
Db 1081 CTGACAGTGGGATATGTCTGCGCAAGCTTTGAAGAAAGCACTTCGCTGCTGCGAG 1140
Qy 3170 GATTACAGGACCTGATTTTGGCCATGATATTTACGACCGGAGGTTGCACTTAAACGTT 3229
Db 1141 GATTACAGGACCTGATTTTGGCCATGATATTTACGACCGGAGGTTGCACTTAAACGTT 1200
Qy 3230 CTGATGCTGCTGCTGCTGCTGATGAGGAGGCTGCGCGAGAGATGATCAACCAAGTTG 3289
Db 1201 CGATGCTGCTGCTGCTGCTGATGAGGAGGCTGCGCGAGAGATGATCAACCAAGTTG 1260
Qy 3290 CACAGAAAGATGAGTGCATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3349
Db 1261 CGAGAAAGATGAGTGCATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Qy 3350 CGATCTACCAAGAGCTCTTCTCCAGAGGAGGAGTGGGGGTTATCATGATCTGTATTAG 3409
Db 1321 CGATCTACCAAGAGCTCTTCTCCAGAGGAGGAGTGGGGGTTATCATGATCTGTATTAG 1380
Qy 3410 GTTTCGAACTGTCCGAGTACATCCGAGGCAATATTTCCTGAGTGGAGTGTGTC 3469
Db 1381 GTTTCGAACTGTCCGAGTACATCCGAGGCAATATTTCCTGAGTGGAGTGTGTC 1440
Qy 3470 GTTGTCTGTGATGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3529
Db 1441 GTTGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
Qy 3530 TGATGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3589
Db 1501 TGATGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
Qy 3590 CTTTCCGGGCTTCAAGAGTGAACACTATTGAGATGATGATGATGATGATGATGATGATG 3649
Db 1561 C-TTTCGACCTTCAAGAGTGAACACTATTGAGATGATGATGATGATGATGATGATG 1619
Qy 3650 TGGCGAATGTACAATATCTTGCAGAAATCCCTTGAT 3686
Db 1620 TGGCGAATGTACAATATCTTGCAGAAATCCCTTGAT 1656

RESULT 6
AAFP1528
ID AAFP1528 standard; DNA; 1527 BP.
XX
AC AAFP1528;
XX
DT 09-APR-2001 (first entry)
XX
DE C-glutamicum phosphoenolpyruvate DNA #1.
XX
KW Phosphoenolpyruvate; sugar phosphotransferase system; PTS; ds.
XX
OS Corynebacterium glutamicum.
XX
W0200102583-A2.
XX
PD 11-JAN-2001.
XX
PF 27-JUN-2000; 2000WO-IB000973.
XX
PR 01-JUL-1999; 99US-0142691P.
PR 23-AUG-1999; 99US-0150310P.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042097.
XX
PA (BAD1 ) BASF AG.
XX
PI Pompejus M, Kroegeer B, Schroeder H, Zelder O, Habernauer G;
DR WPL; 2001-080989/09.
XX
XX Corynebacterium glutamicum nucleic acids encoding phosphoenolpyruvate:
PT sugar phosphotransferase system proteins or their portions, useful for
XX typing or identifying C. glutamicum or related bacteria, and as markers
```

PT for transformation.

XX PS Claim 3; Page 98-101; 144bp; English.

CC The present invention relates to Corynebacterium glutamicum
CC phosphoenolpyruvate: sugar phosphotransferase system (PTS) proteins. The
CC PTS nucleic acids and proteins are useful in the identification of
CC microorganisms which can be used to produce fine chemicals, for
CC modulating fine chemical production in C. glutamicum or related bacteria,
CC the typing or identification of C. glutamicum or related bacteria, as
CC reference points for mapping C. glutamicum genome, and as markers for
CC transformation

XX Sequence 1527 BP; 304 A; 392 C; 430 G; 401 T; 0 U; 0 Other;

Query Match 23.7%; Score 1412.6; DB 5; Length 1527;
Best Local Similarity 95.7%; Pired. No. 0;
Matches 1462; Conservative 0; Mismatches 64; Indels 1; Gaps 1;

QY 4259 CTGATGGCATCTGCGCGTTGCGTCTTGCCAGTGTGGTGGTTTCAACCGAACCAAG 4318
DB 1 CTCATGGCATCTGCGCGTTGCGTCTTGCCAGTGTGGTGGTTTCAACCGAACCAAG 60
QY 4319 CGTTTCGTCGCAATGATGTTCTGCGCGCGG-GCATTGGTATGGCGATGTTCCCAAC 4377
DB 61 CGTTTCGCGGCAATGATGTTCTGCGCGCGGATTTGGTATGGCGATGTTCCCGAG 120
QY 4378 CCTGTGAACGGCTACGACGTGGCGCCACCATGACCGCGGCGAAATGCGCAATGTGTC 4437
DB 121 CTGTGGAAACGGCTACGACGTGGCGCCACCATGACGTGGCGGCGAAATGCGCAATGTGTC 180
QY 4438 CCGTTTGGTGGATGTTGCTCAAGCTGGTTACGAGGCGACCGTCTCTGTGTGAT 4497
DB 181 CCGTTTGGTGGATGTTGCTCAAGCTGGTTACGAGGCGACCGTCTCTGTGTGAT 240
QY 4498 GGTCTCTGGATCTGCGCAACGATCGAAGATTCCTGCAACAAGCGACTATGGGCACTGC 4557
DB 241 GGTCTCTGGATCTGCGCAACGATCGAAGATTCCTGCAACAAGCGACTATGGGCACTGC 300
QY 4558 AAGACTTCTGATCAACCCAGTGTGACTCTGCTGCTACCGCGCTTCTTACGTTCAATGTC 4617
DB 301 AAGACTTCTGATCAACCCAGTGTGACTCTGCTGCTACCGCGCTTCTTACGTTCAATGTC 360
QY 4618 TATTTGGTCAGCAATGCGTGGTGGTGAATCTTGTCTGCGACAACGCTTCTGAGGATCTTA 4677
DB 361 CATTTGCCAGCAATGCGTGGTGGTGAATCTTGTCTGCGACAACGCTTCTGAGGATCTTA 420
QY 4678 TGAATTCGATGTCAGTGGCGGCTCTGCTTTCGCTGTGCTACTCAACCAATGCTAT 4737
DB 421 TGAATTCGATGTCAGTGGCGGCTCTGCTTTCGCTGTGCTACTCAACCAATGCTAT 480
QY 4738 CACTGGTCTGACCAAGTCTCTCCCGCAATTGAGCTGAGCTGTTCACACGAGGTGATC 4797
DB 481 CACTGGTCTGACCAAGTCTCTCCCGCAATTGAGCTGAGCTGTTCACACGAGGTGATC 540
QY 4798 CTTCAATCTTCCGCAACCGCATTCATGCGCAATATCGGCGAGGTGACGATGTTGGCACT 4857
DB 541 CTTCAATCTTCCGCAACCGCATTCATGCGCAATATCGGCGAGGTGACGATGTTGGCACT 600
QY 4858 GTTCTTCCAGCAAGAGTGAAGCTCAAGGCGCTTGAAGGCTTCAAGGTCTCCGCG 4917
DB 601 GTTCTTCCAGCAAGAGTGAAGCTCAAGGCGCTTGAAGGCTTCAAGGTCTCCGCG 660
QY 4918 TGTTCCTGGTATTAACAGACCTGCGATCTTGGGTGAACCTTCCGCTGCGTGGCGCTT 4977
DB 661 TGTTCCTGGTATTAACAGACCTGCGATCTTGGGTGAACCTTCCGCTGCGTGGCGCTT 720
QY 4978 CTACATGGTATCGGATCGGACGCTATGCGTGGCGCTTGAATGCACTTTGAATCA 5037
DB 721 CTACATGGTATCGGATCGGACGCTATGCGTGGCGCTTGAATGCACTTTGAATCA 780
QY 5038 GGCAGTGGCTGGCGCTGAGGTTTCTGGGTGTTGTTTATGATGCTCCAGATAT 5097

DB 781 GGCAGTGGCTGGCGCTGAGGTTTCTGGGTGTTGTTTATGATGCTCCAGATAT 840
QY 5098 GGTATGTTCTTGTGTTGCGCGGTACTTCTTGTATGCAATTCGCGCGAGCATTC 5157
DB 841 GGTATGTTCTTGTGTTGCGCGGTACTTCTTGTATGCAATTCGCGCGAGCATTC 900
QY 5158 TTAATGCTTTTCTTGTGTTGCGCGAGGCGAGCATTCGAGTATGCAACCGCTGTC 5217
DB 901 TTAATGCTTTTCTTGTGTTGCGCGAGGCGAGCATTCGAGTATGCAACCGCTGTC 960
QY 5218 AGTCCCTGAGCAACCAAGCCGAGACAGAGCAGAGCAGGAGATTTTCAACGATTC 5277
DB 961 AGTCCCTGAGCAACCAAGCCGAGACAGAGCAGAGCAGGAGATTTTCAACGATTC 1020
QY 5278 CACCATATCCAGGCACTTTTGAACCGGTATGCACTGAGCAGGCTGAGCATTC 5337
DB 1021 CACCATATCCAGGCACTTTTGAACCGGTATGCACTGAGCAGGCTGAGCATTC 1080
QY 5338 CATGTTTCCAGCGAAGCACTTGGCTGAGTGTGCGATGTCCTCCCAAGGCGAGCT 5397
DB 1081 CATGTTTCCAGCGAAGCACTTGGCTGAGTGTGCGATGTCCTCCCAAGGCGAGCT 1140
QY 5398 GGTTCACAGTGAAGCGAAGATGCTGAGTGTGCTTCCCATGCTGTCAGCTTTCAGT 5457
DB 1141 GGTTCACAGTGAAGCGAAGATGCTGAGTGTGCTTCCCATGCTGTCAGCTTTCAGT 1200
QY 5458 CCGCACTAAGGCTGAGATGTTTCAATGTGATATCTTGAATGCAATTTGTTTGCAC 5517
DB 1201 TCGCACTAAGGCTGAGATGTTTCAATGTGATATCTTGAATGCAATTTGTTTGCAC 1260
QY 5518 CGTAAACCTCAACCGGACGACCTTTTACCGGCTGAAGAGCGGCGAGTCAAAAC 5577
DB 1261 AGTAAACCTCAACCGGACGACCTTTTACCGGCTGAAGAGCGGCGAGTCAAAAC 1320
QY 5578 AGGAGAGCTGCTGTGTAATGATATTGATCCATTAAAGCTGCAAGTTATGAGTAAAC 5637
DB 1321 AGGAGAGCTGCTGTGTAATGATATTGATCCATTAAAGCTGCAAGTTATGAGTAAAC 1380
QY 5638 CAGCGGATGTTGTTTGAATTAACAAGAAACGAGCTGTAAACCTTACGTTTGGG 5697
DB 1381 CAGCGGATGTTGTTTGAATTAACAAGAAACGAGCTGTAAACCTTACGTTTGGG 1440
QY 5698 CGAAATTAAGGAGGAGCAACCTGCTCAACGTCGCAAAAGAAAGCGGTCCAGCAAC 5757
DB 1441 CGAAATTAAGGAGGAGCAACCTGCTCAACGTCGCAAAAGAAAGCGGTCCAGCAAC 1500
QY 5758 ACCATTAAGTGAACCTTGAAGTTCG 5784
DB 1501 ACCATTAAGTGAACCTTGAAGTTCG 1527

RESULT 7
AB865346
ID AB865346 strand; DNA; 1527 BP.
XX
AC AB865346;
XX
DT 15-NOV-2002 (first entry)
XX
DE DNA encoding C. glutamicum metabolic pathway (MP) protein #5.
XX
KW Metabolic pathway protein; MP; cell metabolism; amino acid; vitamin;
KW cofactor; nucleotide; nucleoside; trehalose; fine chemical production;
KW organic acid; non-proteinogenic amino acid; purine base; carbohydrate;
KW pyrimidine base; lipid; unsaturated fatty acid; diols; polyketide;
KW aromatic compound; food industry; animal feed; cosmetic industry;
KW pharmaceutical industry; gene; db.
OS Corynebacterium glutamicum ATCC 13032.
XX
PN WO200251231-A1.
XX
PD 04-JUL-2002.

XX 22-DEC-2000; 2000WO-EP013143.
 XX 22-DEC-2000; 2000WO-EP013143.
 XX (BADI) BASF AG.
 PA Pompejus M, Kroegeer B, Zelder O, Schroeder H;
 XX MPI; 2002-643289/69.
 DR P-PSDB; ABG80325.
 PT New metabolic pathway genes of *Corynebacterium glutamicum* for producing
 PT fine chemicals, e.g. lipids, (un)saturated fatty acids, vitamins,
 PT cofactors or enzymes used in food, feed, cosmetics or pharmaceutical
 PA industries.
 XX
 PS Claim 1; Page 108-111; 176pp; English.
 CC The present invention relates to the isolation of *Corynebacterium*
 CC glutamicum metabolic pathway (MP) proteins, and the polynucleotide
 CC sequences encoding them. The MP proteins are enzymes involved in the
 CC metabolism of molecules important for the normal functioning of cells
 CC (e.g. amino acids, vitamins, cofactors, nucleotides and nucleosides, or
 CC trehalose). The polynucleotide sequences encoding the MP proteins are
 CC useful for producing fine chemicals, particularly organic acids, non-
 CC proteinogenic amino acids, purine and pyrimidine bases, nucleosides,
 CC nucleotides, lipids, (un)saturated fatty acids, diols, carbohydrates,
 CC aromatic compounds, vitamins, cofactors, polyketides and enzymes. The
 CC fine chemicals are useful in the food, animal feed, cosmetic or
 CC pharmaceutical industries. ABS65342-ABS65364 encode the C. glutamicum MP
 CC proteins of the invention
 XX

Sequence 1527 BP; 304 A; 392 C; 430 G; 401 T; 0 U; 0 Other;

Query Match 23.7%; Score 1412.6; DB 6; Length 1527;

Best Local Similarity 95.7%; Pred. No. 0;

Matches 1462; Conservative 0; Mismatches 64; Indels 1; Gaps 1;

QY 4259 CTGATGCACTGCGCGCTTGGGCTTCTGCAAGTGTGGTTTCCAGCCGCAACGA 4318
 DB 1 CTCATGGACCTGCGCGCTTGGGCTTCTGCAAGTGTGGTTTCCAGCCGCAACGA 60
 QY 4319 CGTTTGGTGGCAATAGTTCCTGGGCGCGG-GCATTTGATAGCGATGTGTCCCAAC 4377
 DB 61 CGTTTGGGCGCAATAGTTCCTGGGCGCGGCGATTTGATAGCGATGTGTCCCAAC 120
 QY 4378 CCGTGTAAACGGCTACGACGTGGCCGCAACATGACCGGGGGAATGCCAATGTGTC 4437
 DB 121 CTTGTGGAACGGCTACGACGTGGCCGCAACATGACCGGGGGAATGCCAATGTGTC 180
 QY 4438 CCTGTTGGTGTGATGTGCTCAAGCTGTGTTACAGGGACCGGTCTTGTGTGTGT 4497
 DB 181 CCGTTTGGTGTGATGTGCTCAAGCTGTGTTACAGGGACCGGTCTTGTGTGTGT 240
 QY 4498 GGTCTCTTGGATTTCTGGCAACATTCAGAAAGTTCCTGCAACGCACTAATGGCACTGC 4557
 DB 241 GGTCTCTTGGATTTCTGGCAACATTCAGAAAGTTCCTGCAACGCACTAATGGCACTGC 300
 QY 4558 AGACTCTGATCAACCCAGTGTGCTGCTGCTCAACCGGCTTCTTACGTTCACTTGC 4617
 DB 301 AGACTCTGATCAACCCAGTGTGCTGCTGCTCAACCGGCTTCTTACGTTCACTTGC 360
 QY 4618 TATTTGTCAAGCAATCGCTGGGTGGGTGACTTGTGGCAACGGTCTGCAAGGACTCTA 4677
 DB 361 CATTTGCCAGCAATCGCTGGGTGGGTGACTTGTGGCAACGGTCTGCAAGGACTCTA 420
 QY 4678 TATTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 4737
 DB 421 TATTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 480
 QY 4738 CACTGCTGCAACCACTCTTCCGCAATTGAGCTGAGCTGTTCAACGAGGTGATC 4797

DB 481 CACTGCTGCAACCACTCTTCCGCAATTGAGCTGAGCTGTTTAAACGAGGTGATC 540
 QY 4798 CTTCACTTTGCAACCGCATTCATGCGCAATATGCGGAGGTGACAGATTTTGGCAGT 4857
 DB 541 CTTCACTTTGCAACCGCATTCATGCGCAATATGCGGAGGTGACAGATTTTGGCAGT 600
 QY 4858 GTTCTTCTAGCAAGAGTGAAGAGCTCAAGGGCTTGCAGGTCTCAGGTCTCCGC 4917
 DB 601 GTTCTTCTAGCAAGAGTGAAGAGCTCAAGGGCTTGCAGGTCTCAGGTCTCCGC 660
 QY 4918 TGTCTTGTGATTAACAGAGCTGCGATCTTGTGTGTAACCTTGCCTGCGCTGCGCTT 4977
 DB 661 TGTCTTGTGATTAACAGAGCTGCGATCTTGTGTGTAACCTTGCCTGCGCTGCGCTT 720
 QY 4978 CTCAATTTGATTCGATACCGGATGCTATGCTGCGCTTGTGATGCACTTTGATATCA 5037
 DB 721 CTCAATTTGATTCGATACCGGATGCTATGCTGCGCTTGTGATGCACTTTGATATCA 780
 QY 5038 GGCAGTTGCGGTGGGCGCTGAGAGTTCTTGGGTGTGTTCTATGATGCTCCAGATAT 5097
 DB 781 GGCAGTTGCGGTGGGCGCTGAGAGTTCTTGGGTGTGTTCTATGATGCTCCAGATAT 840
 QY 5098 GGTCAATGTTCTTGTGTTGGCGGTAGTACCTTTGTCAATGCAATTCGCGGAGGATTC 5157
 DB 841 GGTCAATGTTCTTGTGTTGGCGGTAGTACCTTTGTCAATGCAATTCGCGGAGGATTC 900
 QY 5158 TTATGCGCTTTATCTTGTGTTGCGCGGACGCAAGCATTTGATGCAATGCAATTCGCTCC 5217
 DB 901 TTATGCGCTTTATCTTGTGTTGCGCGGACGCAAGCATTTGATGCAATGCAATTCGCTCC 960
 QY 5218 AGTGCCTGAGAAACGACCAAGCGCAAGAGCAAGAGCAACCGCAGATTTCAAAAGATTC 5277
 DB 961 AGTGCCTGAGAAACGACCAAGCGCAAGAGCAAGAGCAACCGCAGATTTCAAAAGATTC 1020
 QY 5278 CACCATCATCCAGGACCTTTGACCGGTGAAGCTATGCACTGAGACGCTGACGATGC 5337
 DB 1021 CACCATCATCCAGGACCTTTGACCGGTGAAGCTATGCACTGAGACGCTGACGATGC 1080
 QY 5338 CATGTTTGGCAGGAAAGCTTGGCTCAGGTGTGCGCATGCTCCCAACCAAGGGGACCT 5397
 DB 1081 CATGTTTGGCAGGAAAGCTTGGCTCAGGTGTGCGCATGCTCCCAACCAAGGGGACCT 1140
 QY 5398 GGTTCACACAGTGAACGGAAGATGCTGTGAGCTTCCATCTGATCAGCTTTGCGAGT 5457
 DB 1141 AGTTTCTCGGTGATGAGAAAGATTTGTGTGCAATCTGCGCATGCTTTGCGAGT 1200
 QY 5458 CCGCATTAAGGCTGAGATGTTCCATGTGATATCTTGAATGCAATTTGTTGCAAC 5517
 DB 1201 CCGCATTAAGGCTGAGATGTTCCATGTGATATCTTGAATGCAATTTGTTGCAAC 1260
 QY 5518 CGTAAACCTCAACGCAAGCACTTTAACCGCTGAAGAAAGAGGCGATGAAGTCAAGC 5577
 DB 1261 AGTAAACCTCAACGCAAGCACTTTAACCGCTGAAGAAAGAGGCGATGAAGTCAAGC 1320
 QY 5578 AGGGAGCTGCTGTGATATGATATGATGATGATGATGATGATGATGATGATGATGATG 5637
 DB 1321 AGGGAGCTGCTGTGATATGATATGATGATGATGATGATGATGATGATGATGATGATG 1380
 QY 5638 CACGCGATGTTGTTTGAATTAACAAGAAACCGGACCTGTAAACCTTACGTTTGGG 5697
 DB 1381 CACGCGATGTTGTTTGAATTAACAAGAAACCGGACCTGTAAACCTTACGTTTGGG 1440
 QY 5698 CGAAATTAAGGAGGAGCAACCTGCTCAAGTGCAGAAAGAAAGAGGAGGAGGAGGAGGAG 5757
 DB 1441 CGAAATTAAGGAGGAGCAACCTGCTCAAGTGCAGAAAGAAAGAGGAGGAGGAGGAGGAG 1500
 QY 5758 ACCATAAGTTGAACCTTGAAGTTCG 5784
 DB 1501 ACCATAAGTTGAACCTTGAAGTTCG 1527

RESULT 8
 AAF71533

ID AAF71533 standard; DNA; 1342 BP.
XX AAF71533;
AC
XX
XX
XX 30-APR-2001 (first entry)
DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:347.
XX
XX
XX Corynebacterium glutamicum; carbon metabolism and energy production;
KM SMP protein; sugar metabolism and oxidative phosphorylation protein;
KM fine chemical production; organic acid; proteinogenic amino acid;
KM nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KM nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KM carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
KM diagnosis; Corynebacterium diptheriae; evolutionary study; ds.
XX
XX Corynebacterium glutamicum.
XX
XX WO200100844-A2.
XX
XX 04-JAN-2001.
PD
XX
XX 23-JUN-2000; 2000WO-IB000943.
PF
XX
XX 25-JUN-1999; 99US-0141031P.
PR 08-JUL-1999; 99DE-01031412.
PR 08-JUL-1999; 99DE-01031412.
PR 08-JUL-1999; 99DE-01031419.
PR 08-JUL-1999; 99DE-01031420.
PR 08-JUL-1999; 99DE-01031424.
PR 08-JUL-1999; 99DE-01031428.
PR 08-JUL-1999; 99DE-01031431.
PR 08-JUL-1999; 99DE-01031433.
PR 08-JUL-1999; 99DE-01031434.
PR 08-JUL-1999; 99DE-01031510.
PR 08-JUL-1999; 99DE-01031562.
PR 08-JUL-1999; 99DE-01031634.
PR 09-JUL-1999; 99DE-01032180.
PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-01032230.
PR 09-JUL-1999; 99US-0143208P.
PR 14-JUL-1999; 99DE-01032924.
PR 14-JUL-1999; 99DE-01032973.
PR 14-JUL-1999; 99DE-01033005.
PR 27-AUG-1999; 99DE-01040765.
PR 31-AUG-1999; 99US-0151572P.
PR 03-SEP-1999; 99DE-01042076.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042086.
PR 03-SEP-1999; 99DE-01042087.
PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042123.
PR 03-SEP-1999; 99DE-01042125.
XX
XX (BADI) BASF AG.
XX
XX
XX Pompejus M, Kroegeer B, Schroeder H, Zelder O, Haberhauer G;
PI WPI; 2001-061975/07.
XX
XX P-PSDB; AAB79416.
DR
XX
XX
XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
PT metabolism and oxidative phosphorylation protein for production or
PT modulation of production of fine chemicals e.g. amino acids,
XX carbohydrates or enzymes.
XX
XX Claim 3; Page 627-629; 1246pp; English.
XX
XX AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
CC metabolism and oxidative phosphorylation (SMP) proteins given in
CC to AAB 79633 which are involved in carbon metabolism and energy
CC production. The C. glutamicum SMP gene can be used in vectors (II) for

CC expression in host cells and production or modulation of production of
CC fine chemicals, such as, an organic acid, a proteinogenic or
CC nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a
CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
CC polyketide, or an enzyme. The presence of (I) or SMP proteins (III)
CC encoded by them are used for diagnosing the presence or activity of
CC Corynebacterium diptheriae in a subject. (I), (II), (III) or host cells
CC containing them are used to map genomes of organisms related to C.
CC glutamicum, identify and localise C. glutamicum sequences of interest, in
CC evolutionary studies, in determining SMP protein regions required for
CC function, in modulating SMP protein activity, in modulating the
CC metabolism of sugars, and in modulating high-energy molecule production
CC in a cell (i.e. xrp, NADPH)
XX
XX
SQ Sequence 1342 BP; 312 A; 373 C; 359 G; 298 T; 0 U; 0 Other;
XX
XX
Query Match 22.1%; Score 1321.2; DB 4; Length 1342;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1329; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 2238 ATGAAGACAGATATCCAAAGCTGAAAAAGCTGATCTACCGTCTCATGAGCAATTAA 2297
DB 1 ATGAAGACAGATATCCAAAGCTGAAAAAGCTGATCTACCGTCTCATGAGCAATTAA 60
QY 2298 AGCTGGCTGAGAAACAAAGAAAGTACTGTGGGGCTATGACACAGAACTTTCAG 2357
DB 61 AGCTGGCTGAGAAACAAAGAAAGTACTGTGGGGCTATGACACAGAACTTTCAG 120
QY 2358 TTTGGGCGCTGCGTACATGTACTCTCCGAGGGAGCTCAATGATCCCAAGGAT 2417
DB 121 TTTGGGCGCTGCGTACATGTACTCTCCGAGGGAGCTCAATGATCCCAAGGAT 180
QY 2418 GTACGTGATGAGATACCTTCACGCTCTACTACAGACGATCCAGGTTCCCTTCCG 2477
DB 181 GTACGTGATGAGATACCTTCACGCTCTACTACAGACGATCCAGGTTCCCTTCCG 240
QY 2478 ACCAAAGCGACCGGCTGGGGCTCACACACACGCGCTTGAACCGGACCGCAATTGA 2537
DB 241 ACCAAAGCGACCGGCTGGGGCTCACACACACGCGCTTGAACCGGACCGCAATTGA 300
QY 2538 GTGACGACACCTGCGCGACGCTCTTAAACCGGATGATCTATGACCTGATGATGCTA 2597
DB 301 GTGACGACACCTGCGCGACGCTCTTAAACCGGATGATCTATGACCTGATGATGCTA 360
QY 2598 TTCGGGTGAGCGGATTTACTGACGACCACTTAACTTTCTACACCGGCAACTTAA 2657
DB 361 TTCGGGTGAGCGGATTTACTGACGACCACTTAACTTTCTACACCGGCAACTTAA 420
QY 2658 AATTGACGAAAGCGCGCGGACCAACCAAACTTGTCAAGTCGAGAACCACTGGGCT 2717
DB 421 AATTGACGAAAGCGCGCGGACCAACCAAACTTGTCAAGTCGAGAACCACTGGGCT 480
QY 2718 GATGGGCGGCAATTCATCGCGCTTGAATCCGCTTATCGACGACCGCGGCGG 2777
DB 481 GATGGGCGGCAATTCATCGCGCTTGAATCCGCTTATCGACGACCGCGGCGG 540
QY 2778 TTTCACACCCCATTAACCGGATCCCATGATCAGCCCTGATGATGATGATGATGAT 2837
DB 541 TTTCACACCCCATTAACCGGATCCCATGATCAGCCCTGATGATGATGATGATGAT 600
QY 2838 TCTTGGGGCCCAAGCGGAAACCTTCAACCGGTGACAGGCTTCTATACCGTGGACGAT 2897
DB 601 TCTTGGGGCCCAAGCGGAAACCTTCAACCGGTGACAGGCTTCTATACCGTGGACGAT 660
QY 2898 TGAAGAACTGGGAATTTCTCGGATGAATCACTTTGACCTCAGTGAACCACTGGTTC 2957
DB 661 TGAAGAACTGGGAATTTCTCGGATGAATCACTTTGACCTCAGTGAACCACTGGTTC 720
QY 2958 TGCCTCTGATCTGTTCCCGATGAGGTACATGATGAGAAATGCCCAACTTTTACGCTTCG 3017
DB 721 TGCCTCTGATCTGTTCCCGATGAGGTACATGATGAGAAATGCCCAACTTTTACGCTTCG 780

OY	3018	CGATGAAGAAA	CTGGCCGAAGATCTG	AGTGCTGATTTTCTGTC	CAAGAAGTTGACCG	3077
Db	781	CGATGAAGAAA	CTGGCCGAAGATCTG	AGTGCTGATTTTCTGTC	CAAGAAGTTGACCG	840
OY	3078	AATCCAGATGA	GGTTACTCACTA	CGAAGCTCTGAC	CCAGTGCAGATATGTCGACAA	3137
Db	841	AATCCAGATGA	GGTTACTCACTA	CGAAGCTCTGAC	CCAGTGCAGATATGTCGACAA	900
OY	3138	GCTTGAAGAA	CGACCTTCCG	GCTTTCGAGAGATT	CAGCCAGCTGATTTCCG	3197
Db	901	GCTTGAAGAA	CGACCTTCCG	GCTTTCGAGAGATT	CAGCCAGCTGATTTCCG	960
OY	3198	ATTTCAGCA	CCGACAGTTG	CGATTAACGGTTCTG	ATGCGTGGCTCGAGATTCG	3257
Db	961	ATTTCAGCA	CCGACAGTTG	CGATTAACGGTTCTG	ATGCGTGGCTCGAGATTCG	1022
OY	3258	GCTGCCCGC	GACGATGATCA	CCCAACAGTTG	CACAGAAAGATGGTGAC	3317
Db	1021	GCTGCCCGC	GACGATGATCA	CCCAACAGTTG	CACAGAAAGATGGTGAC	1080
OY	3318	TGTGCCCCG	CAAGCTTCATT	TGCGCAACG	CGATCTTCTTCC	3377
Db	1081	TGTGCCCCG	CAAGCTTCATT	TGCGCAACG	CGATCTTCTTCC	1140
OY	3378	GGGGGAGT	CGGGGGATACAG	ATCTGTATTAAGT	CTGAACCTGACGAGTAC	3437
Db	1141	GGGGGAGT	CGGGGGATACAG	ATCTGTATTAAGT	CTGAACCTGACGAGTAC	1200
OY	3438	AGGCATATTT	CCCTCGAGTGGAT	GTGTCGTTGTCT	GTGATCTGTATGTGATCG	3497
Db	1201	AGGCATATTT	CCCTCGAGTGGAT	GTGTCGTTGTCT	GTGATCTGTATGTGATCG	1260
OY	3498	TGCGCTAGCT	GATGAATAACT	GTGGCGAATT	TAATGATCGCGACATATTA	3557
Db	1261	TGCGCTAGCT	GATGAATAACT	GTGGCGAATT	TAATGATCGCGACATATTA	1320
OY	3558	GATAAC	TGCGAGTATGAC	AG 3579		
Db	1321	GATAAC	TGCGAGTATGAC	AG 1342		

```

RESULT 9
AAH67870
ID AAH67870 standard; DNA; 1299 BP.
AC AAH67870;
XX
XX 26-SEP-2001 (first entry)
XX
XX C glutamicum coding sequence fragment SEQ ID NO: 2905.
XX
XX Corynebform bacterium; amino acid synthesis; vitamin; saccharide;
XX organic acid synthesis; de.
XX
XX Corynebacterium glutamicum.
OS
XX
XX EP1108790-A2.
XX
XX 20-JUN-2001.
XX
XX 18-DEC-2000; 2000EP-00127688.
XX
XX 16-DEC-1999; 99JP-00377484.
XX
XX 07-APR-2000; 2000JP-00159162.
XX
XX 03-AUG-2000; 2000JP-00280988.
XX
XX (KYOW ) KYOWA HAKKO KOGYO KK.
XX
XX Nakagawa S, Mitsuuchi H, Ando S, Hayashi M, Ochiai K, Yokoi H,
XX Tateishi N, Senoh A, Ikeda M, Ozaki A;
XX
XX WPI; 2001-376931/40.
XX
XX P-PSDB; AAG92651.
XX

```

XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying
PT mutation point of a gene, measuring expression of a gene, analyzing
PT expression profile or pattern of a gene and identifying homologous gene.
XX
XX
XX Claim 8, SEQ ID NO 2905; 246bp + Sequence Listing; English.
PS
XX
CC The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium bacterium, measuring expression amount and analyzing
CC the expression profile or expression pattern of a gene derived from
CC Corynebacterium bacterium, and identifying a homologue of a gene derived from
CC Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid, described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX
XX
SQ Sequence 1299 BP; 287 A; 363 C; 350 G; 299 T; 0 U; 0 Other;

Query Match	21.2%	Score 1267.8	DB 5	Length 1299	
Best Local Similarity	99.0%	Pred. No. 0			
Matches 1286	Conservative	0	Mismatches 12	Indels 1	Gaps 1
QY	2337	TGTGTGGGGCTATGACACAGAACTTTCAGATTGCGCCCTGCGTACATGTGACTCTC	2386		
DB	2	TGTGTGGGGCTATGACACAGAACTTTCAGATTGCGCCCTGCGTACATGTGACTCTC	61		
QY	2387	CGCAGGGGAGGCTCAATGATCCCAACGGAAATGTAAGTACATGATACCTTCACGCT	2446		
DB	62	CGCAGGGGAGGCTCAATGATCCCAACGGAAATGTAAGTACATGATACCTTCACGCT	121		
QY	2447	ACTACAGACAGATCCAGGTTTCCCTTCGACCAAAACGACCGGCTGAGCTCACACA	2507		
DB	122	ACTACAGACAGATCCAGGTTTCCCTTCGACCAAAACGACCGGCTGAGCTCACACA	181		
QY	2507	CCACGCCCTGTACCGGACCCGACGGATTGACGTGACCGACCTGCGCAGCTCTTTAC	2566		
DB	182	CCACGCCCTGTACCGGACCCGACGGATTGACGTGACCGACCTGCTTTAC	241		
QY	2567	CGATGCACTCTATGACCTGATGATGCTATTCGGGTGAGCCGTATTATCTGACGGCA	2622		
DB	242	CGATGCACTCTATGACCTGATGATGCTATTCGGGTGAGCCGTATTATCTGACGGCA	301		
QY	2627	CACCTAAACTTTTCTCAACACGGGCAACCTAAAAATTTAGAGAAAGCGCGCCACCCAAA	2686		
DB	302	CACCTAAACTTTTCTCAACACGGGCAACCTAAAAATTTAGAGAGCGCGCCACCCAAA	361		
QY	2687	ACCTTGTGGAAGTCGAGGACCCAACTGCGGCTGATGAGCGGCAATTCATGCGCGTTC	2746		
DB	362	ACCTGTCGAAGTCGAGGACCCAACTGCGGCTGATGAGCGGCAATTCATGCGCGTTC	421		
QY	2747	AAATCCGCTTATCGACGGACCCCGCAGCGTTTCAACCCCATTTACCGCGATCCCATGA	2806		
DB	422	AAATCCGCTTATCGACGGACCCCGCAGCGTTTCAACCCCATTTACCGCATTCACATGA	481		
QY	2807	TCAGGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2866		
DB	482	TCAGGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	541		
QY	2867	GTCGACGCGGTTCTATACCGCTCGACAGATCTTGGAAATCTGCGGTGAATATCA	2922		
DB	542	GTCGACGCGGTTCTATACCGCTCGACAGATCTTGGAAATCTGCGGTGAATATCA	601		
QY	2927	CCTTTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2986		
DB	602	CCTTTGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	661		
QY	2987	TGTGGGAATGCCCCAACCTTTTTCAGTTGCGGTGCGATGAAAGAAATGCGGAATCTCGAG	3046		
DB	662	TGTGGGAATGCCCCAACCTTTTTCAGTTGCGGTGCGATGAAAGAAATGCGGAATCTCGAG	721		


```

QY 362 TCAAGAAAATGAGTCAAGTCAAGTTAAAAATTTGAGGAGAGATTTGTTCCCTCCCTGG 421
DB 121 TCAAGAAAATGAGTCAAGTCAAGTTAAAAATTTGAGGAGAGATTTGTTCCCTCCCTGG 180
QY 422 GGTGATTTGAGTCTTTCTCAACTGAAAAAGGATCATCAAGAACTCTGTGAGAAAC 481
DB 181 GGTGATTTGAGTCTTTCTCAACTGAAAAAGGATCATCAAGAACTCTGTGAGAAAC 240
QY 482 AGCAGTTAAAAAGCAGAGATTCAACCCGAACTCCCAAGATTTGTTCCCTGTTTATTGA 541
DB 241 AGCAGTTAAAAAGCAGAGATTCAACCCGAACTCCCAAGATTTGTTCCCTGTTTATTGA 300
QY 542 TCTTATATATCAAGGAGAAAGGTGGGCGCTTCTTACGAGGAGAGAGAGAGAGAGAG 601
DB 301 TCTTATATATCAAGGAGAAAGGTGGGCGCTTCTTACGAGGAGAGAGAGAGAGAGAG 360
QY 602 GAACAACCGGCGAGTATCAACCGGAAATGAGCAGACCGTATGTTGCCAAGCATGTTTC 661
DB 361 GAATGCGCGGAGTATCAACCGGAAATGAGCAGACCGTATGTTGCCAAGCATGTTTC 420
QY 662 GCGCGCGGCTGAGCGCATGAGCGGCGGAGTGAATACTTATCCCTGTGTGAAGAGT 721
DB 421 GCGCGCGGCTGAGCGCATGAGCGGCGGAGTGAATACTTATCCCTGTGTGAAGAGT 480
QY 722 CCTGCTGTGCGGCAATTCACTGAGGAGCGCTTCTTATCAACGAGTGCCTGTGTGATCA 781
DB 481 CCTGCTGTGCGGCAATTCACTGAGGAGCGCTTCTTATCAACGAGTGCCTGTGTGATCA 540
QY 782 AAACCCGAGATTTCAATTTTCCCGGAGACCAACAGATCTTCCCGGATGATCAATGCGAG 841
DB 541 AAACCCGAGATTTCAATTTTCCCGGAGACCAACAGATCTTCCCGGATGATCAATGCGAG 600
QY 842 AAAAGTTGATCAATGATCACTAGTAGCGCGGAACTGACATTTTCTGACCTTCT 901
DB 601 AAAAGTTGATCAATGATCACTAGTAGCGCGGAACTGACATTTTCTGACCTTCT 660
QY 902 CGATCTCTGCGGAGGAGCAGCATATGCTCTCTTCCGAGCAGTATGAGATTTTGA 961
DB 661 CGATCTCTGCGGAGGAGCAGCATATGCTCTCTTCCGAGCAGTATGAGATTTTGA 720
QY 962 TACCACTACGAGCGAATTTGCTTGGCTAAGAGAAAAATGTCAGCGTCAACGCGTACCA 1021
DB 721 TACCACTACGAGCGAATTTGCTTGGCTAAGAGAAAAATGTCAGCGTCAACGCGTACCA 780
QY 1022 TTTGTTCAATGAGTATGCTCTGCTGATCACTAGGAGTCTCCCGGAGCGGAGCTTGTCT 1081
DB 781 TTTGTTCAATGAGTATGCTCTGCTGATCACTAGGAGTCTCCCGGAGCGGAGCTTGTCT 840
QY 1082 TGCTGCGGACGTCGCGGAGAGCATATGTTGAGTTCGCGAGCGGCGTGCATTTGGC 1141
DB 841 TGCTGCGGACGTCGCGGAGAGCATATGTTGAGTTCGCGAGCGGCGTGCATTTGGC 900
QY 1142 CGATGGAACGCTGATCTAGCTGTTTCAACAGCGCTTTTTCATCAAGGAGCATATGGA 1201
DB 901 CGATGGAACGCTGATCTAGCTGTTTCAACAGCGCTTTTTCATCAAGGAGCATATGGA 960
QY 1202 AGCGCGCGGAATGCGAGAGGAGTATGATTTTGGGCGTTTGAACGTCACCGTACCGGA 1261
DB 961 AGCGCGCGGAATGCGAGAGGAGTATGATTTTGGGCGTTTGAACGTCACCGTACCGGA 1020
QY 1262 TGAAGTTCGCGCTGCTGCGAGTGGCGGCGCATCGCGGAGGAGCAGACATTAAGCAG 1321
DB 1021 TGAAGTTCGCGCTGCTGCGAGTGGCGGCGCATCGCGGAGGAGTACAGACATTAAGCAG 1080
QY 1322 TCAAGTTGTCGACCACTGTCGAGGAGGATGACGCTTATCAACGCGACCTTCACACCTC 1381
DB 1081 TCAAGTTGTCGACCACTGTCGAGGAGGATGACGCTTATCAACGCGACCTTCACACCTC 1140
QY 1382 AACCGTGGCGGCTAAATTTCTGCTTGGAGTATGAGGAGTATGAGGAGTATGAGGAGTAT 1441
DB 1141 AACCGTGGCGGCTAAATTTCTGAGTATGAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTAT 1200
QY 1442 AAATTTTGTGCTTTGACTCAACGAGCGAGTGAAGGATCAATTAAGTATCAAGT 1501

```

```

DB 1201 AAATTTTGTGCTTTGACTCAACGAGGAGTATGAGGAGTATGAGGAGTATGAGGAGTAT 1260
QY 1502 ACTTAAGTACGAGTAATAACTATCTGT 1528
DB 1261 AATTTAATACGAGCAAACTTCTGT 1287

RESULT 11
AAAF71566
ID AAF71566 standard; DNA; 1287 BP.
XX
AC AAF71566;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum SMP protein nucleotide sequence SEQ ID NO:413.
XX
KW Corynebacterium glutamicum; carbon metabolism and energy production;
KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
KW fine chemical production; organic acid; proteinogenic amino acid;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
KW carbohydrate; aromatic compound; vitamin; cofactor; polysaccharide; enzyme;
KW diagnosis; Corynebacterium diptheriae; evolutionary study; ds.
XX
OS Corynebacterium glutamicum.
XX
PN WO200100844-A2.
XX
PD 04-JAN-2001.
XX
PF 23-JUN-2000; 2000WO-IB000943.
XX
PR 25-JUN-1999; 99US-0141031P.
PR 08-JUL-1999; 99DE-01031412.
PR 08-JUL-1999; 99DE-01031413.
PR 08-JUL-1999; 99DE-01031419.
PR 08-JUL-1999; 99DE-01031420.
PR 08-JUL-1999; 99DE-01031424.
PR 08-JUL-1999; 99DE-01031428.
PR 08-JUL-1999; 99DE-01031431.
PR 08-JUL-1999; 99DE-01031433.
PR 08-JUL-1999; 99DE-01031434.
PR 08-JUL-1999; 99DE-01031510.
PR 08-JUL-1999; 99DE-01031562.
PR 08-JUL-1999; 99DE-01031634.
PR 09-JUL-1999; 99DE-01032180.
PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-01032230.
PR 09-JUL-1999; 99DE-0143208P.
PR 14-JUL-1999; 99DE-01032924.
PR 14-JUL-1999; 99DE-01032973.
PR 14-JUL-1999; 99DE-01033005.
PR 27-AUG-1999; 99DE-01040765.
PR 31-AUG-1999; 99US-0151572P.
PR 03-SEP-1999; 99DE-01042076.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042086.
PR 03-SEP-1999; 99DE-01042087.
PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042123.
PR 03-SEP-1999; 99DE-01042125.
XX
PA (BADI ) BASF AG.
XX
PI Pompejus M, Kroegeer B, Schroeder H, Zeider O, Habernauer G;
XX WPI; 2001-061975/07.
XX DR P-PSDB; AAB79449.
XX
XX New isolated Corynebacterium glutamicum nucleic acid encoding a sugar

```

842 AAAAGTTGGATCAATCGATCACAGTAGCCCGGAAC TGACAATCTTCTGAGCTTCT 901

[illegible]

03-AUG-2000; 2000JP-00280988.
PR
XX

AA
PA (KYOW) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tateishi N, Senoh A, Ikeda M, Ozaki A;
 DR WPI; 2001-376931/40.
 DR P-PSDB; AM92653.
 PT Novel polynucleotides derived from *Corynebacterium* bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 8; SEQ ID NO 2907; 246pp + Sequence Listing; English.
 CC The present invention provides a number of nucleotide and protein
 CC sequences from the *Corynebacterium* *Corynebacterium glutamicum*. These
 CC are useful for identifying the mutation point of a gene derived from a
 CC mutant of *Corynebacterium* *Corynebacterium*, measuring expression amount and analysing
 CC the expression profile or expression pattern of a gene derived from
 CC *Corynebacterium* *Corynebacterium*, and identifying a homologue of a gene derived from
 CC *Corynebacterium* *Corynebacterium* bacteria are useful for producing amino
 CC acids, nucleic acids, vitamins, saccharides and organic acids,
 CC particularly L-lysine. The present sequence is a nucleic acid described
 CC in the exemplification of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the European Patent Office
 XX
 SQ Sequence 1152 BP; 273 A; 320 C; 299 G; 260 T; 0 U; 0 Other;
 Query Match 18.6%; Score 1110.4; DB 5; Length 1152;
 Best Local Similarity 97.7%; Pred. No. 0;
 Matches 1126; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
 QY 354 GGGCATTTCAAGAAATGACAGTCAAGGTTAAATAGAGGAAGATTGTC 413
 DB 1 GGGCATTTCAAGAAATGACAGTCAAGGTTAAATAGAGGAAGATTGTC 60
 QY 414 CCCCTCGGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 473
 DB 61 CCCCAAGGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 QY 474 GAGAACACGACACCTTAAATGACAGTCAAGGTTAAATAGAGGAAGATTGTC 533
 DB 121 GAGAACACGACACCTTAAATGACAGTCAAGGTTAAATAGAGGAAGATTGTC 180
 QY 534 TTTATGATCTTCAATATGACAGTCAAGGTTAAATAGAGGAAGATTGTC 593
 DB 181 TTTATGATCTTCAATATGACAGTCAAGGTTAAATAGAGGAAGATTGTC 240
 QY 594 CAGGCGAGAACACGACGATGATGACAGGTTAAATAGAGGAAGATTGTC 653
 DB 241 CAGGCGAGAACACGACGATGATGACAGGTTAAATAGAGGAAGATTGTC 300
 QY 654 ATGTTTGGGCGCGGCTGACAGTCAAGGTTAAATAGAGGAAGATTGTC 713
 DB 301 ATGTTTGGGCGCGGCTGACAGTCAAGGTTAAATAGAGGAAGATTGTC 360
 QY 714 GAAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 773
 DB 361 GAAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 QY 774 GGTGCTCAAAACCCGAGATTTTCCCGGCAACCCAGATCTTCCCGGCTGATC 833
 DB 421 GGTGCTCAAAACCCGAGATTTTCCCGGCAACCCAGATCTTCCCGGCTGATC 480
 QY 834 CATGCGGGAAGAGTTGATCAATGATGACAGTCAAGGTTAAATAGAGGAAGATTGTC 893
 DB 481 CATGCGGGAAGAGTTGATCAATGATGACAGTCAAGGTTAAATAGAGGAAGATTGTC 540
 QY 894 GAGCTTCTGATCTGCGGCGAGGACGACCATGCTGCTGCTGCTGCTGCTGCTGCT 953
 DB 541 GAGCTTCTGATCTGCGGCGAGGACGACCATGCTGCTGCTGCTGCTGCTGCTGCT 600
 QY 954 GATTTTGATACCACTACAGCGCAATGCTTGGCTTAAAGAGAAATGTGACGCTACG 1013

DB 601 GATTTTGATACCACTACAGCGCAATGCTTGGCTTAAAGAGAAATGTGACGCTACG 660
 QY 1014 GCTACGATTTTGTTCATAGCGATGCTTCCGCTGATCATAGAGGCTCCGCGAGCTGGGC 1073
 DB 661 GCTACGATTTTGTTCATAGCGATGCTTCCGCTGATCATAGAGGCTCCGCGAGCTGGGC 720
 QY 1074 GCTTGTGCTGCTGCGGCGAGTGGCGGGAGCGCATATGTTGATGATGCGGAGCGGCTG 1133
 DB 721 GCTTGTGCTGCTGCGGCGAGTGGCGGGAGCGCATATGTTGATGATGCGGAGCGGCTG 780
 QY 1134 CATTTGCGGATGGAACGCGTGAATCTAGCTGTTTCAACAAACGCTTTTTCATCAGGAC 1193
 DB 781 CATTTGCGGATGGAACGCGTGAATCTAGCTGTTTCAACAAACGCTTTTTCATCAGGAC 840
 QY 1194 GCCATGGAAGCCCGCGGAATGCGACAGTGAATCATTTTGGGAGTTTGAACGTAC 1253
 DB 841 GCCATGGAAGCCCGCGGAATGCGACAGTGAATCATTTTGGGAGTTTGAACGTAC 900
 QY 1254 GTCACCGATGAGTGGCCGCTGCGGAGTGGCGGCGCATGCGCGGGGACACGACACA 1313
 DB 901 GTCACCGATGAGTGGCCGCTGCGGAGTGGCGGCGCATGCGCGGGGACACGACACA 960
 QY 1314 CTAGGAGTCAATGCTGTCACACAGTGGCGAGGGGTATGACGCTTATGACGCGACCTG 1373
 DB 961 CTAGGAGTCAATGCTGTCACACAGTGGCGAGGGGTATGACGCTTATGACGCGACCTG 1020
 QY 1374 CACACTTAAACCGTGGCGCTTAAATTTCCGCTCTTGGCGATCAAGAAATGCTTAAATCC 1433
 DB 1021 CACACTTAAACCGTGGCGCTTAAATTTCCGCTCTTGGCGATCAAGAAATGCTTAAATCC 1080
 QY 1434 AACCTGCAATTTTGTGCTTGTGATCACTCAACGCGCAGGTGCAAAAGTCCATTAGT 1493
 DB 1081 AACCTGCAATTTTGTGCTTGTGATCACTCAACGCGCAGGTGCAAAAGTCCATTAGT 1140
 QY 1494 CATCAAGTACTT 1505
 DB 1141 CATCAAGTACTT 1152
 RESULT 13
 AAF31529
 ID AAF31529 standard; DNA; 1109 BP.
 XX
 AC AAF31529;
 XX
 DT 09-APR-2001 (first entry)
 XX
 DE C-glutamicum phosphoenolpyruvate DNA #2.
 XX
 KW Phosphoenolpyruvate; sugar phosphotransferase system; PTS; ds.
 OS
 XX *Corynebacterium glutamicum*.
 PN WO200102583-A2.
 XX
 PD 11-JAN-2001.
 XX
 PF 27-JUN-2000; 2000WO-IB000973.
 XX
 PR 01-JUL-1999; 99US-0142691P.
 PR 23-AUG-1999; 99US-0150310P.
 PR 03-SEP-1999; 99DE-01042095.
 PR 03-SEP-1999; 99DE-01042097.
 XX
 PA (BADI) BASF AG.
 XX
 PI Pompejus M, Kroege B, Schroeder H, Zelder O, Habernauer G;
 XX WPI; 2001-080989/09.
 DR
 XX
 PT *Corynebacterium glutamicum* nucleic acids encoding phosphoenolpyruvate:
 PT sugar phosphotransferase system proteins or their portions, useful for

PR 03-SEP-1999; 99DE-01042079.
 PR 03-SEP-1999; 99DE-01042086.
 PR 03-SEP-1999; 99DE-01042087.
 PR 03-SEP-1999; 99DE-01042088.
 PR 03-SEP-1999; 99DE-01042095.
 PR 03-SEP-1999; 99DE-01042123.
 PR 03-SEP-1999; 99DE-01042125.
 PA (BADI) BASF AG.
 PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
 XX WPI; 2001-061975/07.
 DR P-PSDB; AAB79454.
 PT New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
 PT metabolism and oxidative phosphorylation protein for production or
 PT modulation of production of fine chemicals e.g. amino acids,
 PT carbohydrates or enzymes.
 XX
 PS Claim 3; Page 751-752; 1246pp; English.
 XX
 CC AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
 CC metabolism and oxidative phosphorylation (SMP) proteins given in AAB79243
 CC to AAB 79633 which are involved in carbon metabolism and energy
 CC production. The C. glutamicum SMP gene can be used in vectors (ii) for
 CC expression in host cells and production or modulation of production of
 CC fine chemicals, such as, an organic acid, a proteinogenic or
 CC nonproteinogenic amino acid (preferred), a purine or pyrimidine base, a
 CC nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty acid,
 CC a diol, a carbohydrate, an aromatic compound, a vitamin, a cofactor, a
 CC polyketide, or an enzyme. The presence of (i) or SMP proteins (iii)
 CC encoded by them are used for diagnosing the presence or activity of
 CC Corynebacterium diptheriae in a subject. (i), (ii), (iii) or host cells
 CC containing them are used to map genomes of organisms related to C.
 CC glutamicum, identify and localise C. glutamicum sequences of interest, in
 CC evolutionary studies, in determining SMP protein regions required for
 CC function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. ATP, NADPH)
 XX
 SQ Sequence 882 BP; 259 A; 232 C; 213 G; 178 T; 0 U; 0 Other;
 XX
 Query Match 12.8%; Score 762; DB 4; Length 882;
 Best Local Similarity 93.0%; Pred. No. 1.4e-218;
 Matches 820; Conservative 0; Mismatches 60; Indels 2; Gaps 2;
 QY 1449 GTGGTCTTTGACTCAACCGCCAGGTGCAAAAGTCCATTAGTCATCACTACTTTAA 1508
 DB 1 GTGGTCTTTGACTCAACCGCCAGGTGCAAAAGTCCATTAGTCATCACTACTTTAA 60
 QY 1509 GTACGAGTAAACTATCTCTGAT-TTTAAAGAGTCCACACATGAAATCATCTGCA 1567
 DB 61 ATACGAGCAAACTTCTCTGATATAAAGAGTCCGACCATGACATCATCTATCTGCA 120
 QY 1568 AGACGAGCAAAAGTCCGCAAAAGTTCAGTCCATTCGCAACCTTCCGCAAAAGG 1627
 DB 121 AGACGAGCAAAAGTCCGCAAAAGTTCAGTCCATTCGCAACCTTCCGCAAAAGG 180
 QY 1628 TGGAACTTTGGGGCTTGCACAGAGTCTTACACATGAGTACTTCAAGAGCTGATCG 1687
 DB 181 CGGAACCTTTGGGGCTTGCACATGATCTGACCTTTGAGCACTCAAGAGCTGATCG 240
 QY 1688 CATGATATGAAGTGGGGAGAGTCTATTCAGAAATCGCAAGGATCTTGTGGATGATA 1747
 DB 241 CATGATATGAAGTGGGGAGAGTCTATTCAGAAATCGCAAGGATCTTGTGGATGATA 300
 QY 1748 CGTGGGACCTAACCCGAGCATGAAAGAGCTACTTTAAACATTCGCAAGAGTTCAC 1807
 DB 301 CGTGGGATTTACCGCGAGCATGAAAGAGCTACTTTCAAAACATTCGTAAGAGTTCAC 360
 QY 1808 TGACCAATCGACATCGTTGATGAAGAGGCTTACAGCCAGATGATGCAAACTGATCC 1867

DB 361 TGACCAATCGACATCGTTGATGAAGAGGCTTACAGCCCAATGATGCAAACTGATCC 420
 QY 1868 ATACGAGCAAGTCCGAGAGTATGAGCAAAAGATCGCTCAGAAATCGTTGAAGTTCAA 1927
 DB 421 ATACGAGCAAGTCCGAGAGTATGAGCAAAAGATCGCTCAGAAATCGTTGAAGTTCAA 480
 QY 1928 CTTTGGCATTCGGCGGAAAGG-CACATGCTTTCAATTAACATCATCTTCTGTGAGG 1986
 DB 481 CTTTGGCATTCGGCGGAAAGG-CACATGCTTTCAATTAACATCATCTTCTGTGAGG 540
 QY 1987 ACTGACAAAGTCCAGGCGCTGACCCCTTAAACTGTGAGAGCAACGCTCGATTTCTGA 2046
 DB 541 ACTGACAAAGTCCAGGCGCTGACCCCTTAAACTGTGAGAGCAACGCTCGATTTCTGA 600
 QY 2047 CACCATGCAAGAGTCCCAACCCAGCCGTCAACCCAGGCTTTGGCACTTTGTCCCGGC 2106
 DB 601 CACCATGCAAGAGTCCCAACCCAGCCGTCAACCCAGGCTTTGGCACTTTGTCCCGGC 660
 QY 2107 GCAAAACATCGTGTGGGCAACTGTGAGAGGAAAGCGAGCCGATCCGCGGAACGT 2166
 DB 661 GCAAAACATCGTGTGGGCAACTGTGAGAGGAAAGCGAGCCGATCCGCGGAACGT 720
 QY 2167 GGAAGGCCAGTGAAGTCTTCTTGGCCAGTTTCCATCTGTAGATGCAACAATGCCACC 2226
 DB 721 GGAAGGCCAGTGAAGTCTTCTTGGCCAGTTTCCATCTGTAGATGCAACAATGCCACC 780
 QY 2227 ATCATCGTTGATGAGAGGAGTATCCAGCTGGAAGGCTGTATCTTACCTGTCAT 2286
 DB 781 CATCATCGTTGATGAGAGGAGTATCCAGCTGGAAGGCTGTATCTTACCTGTCAT 840
 QY 2287 GGAGCAATTAAGCTGGCTGAGAAACAAAGGAAAGGAAAGTACTG 2328
 DB 841 GGAGCAATTAAGCTGGCTGAGAAACAAAGGAAAGGAAAGTACTG 882
 XX
 RESULT 15
 AAB67871
 ID AAB67871 standard; DNA; 759 BP.
 XX
 AC AAB67871;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE C glutamicum coding sequence fragment SEQ ID NO: 2906.
 XX
 KW Corynebacterium; amino acid synthesis; vitamin; saccharide;
 KW organic acid synthesis; ds.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN EP108790-A2.
 XX
 PD 20-JUN-2001.
 XX
 PE 18-DEC-2000; 2000EP-00127688.
 XX
 PR 16-DEC-1999; 99JP-00377484.
 PR 07-APR-2000; 2000JP-00159162.
 PR 03-AUG-2000; 2000JP-00280988.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;
 PI Tabeiishi N, Senoh A, Ikeda M, Ozaki A;
 DR WPI; 2001-376931/40.
 DR P-PSDB; AAG92652.
 XX
 PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying
 PT mutation point of a gene, measuring expression of a gene, analyzing
 PT expression profile or pattern of a gene and identifying homologous gene.
 XX
 PS Claim 8; SEQ ID NO 2906; 246pp + Sequence Listing; English.

Search completed: March 9, 2005, 21:00:56
Job time : 2750 secs

XX The present invention provides a number of nucleotide and protein
CC sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These
CC are useful for identifying the mutation point of a gene derived from a
CC mutant of corynebacterium bacterium, measuring expression amount and analysing
CC the expression profile or expression pattern of a gene derived from
CC Corynebacterium bacterium. Corynebacterium bacterium is a gene derived from
CC Corynebacterium bacterium. Corynebacterium bacterium are useful for producing amino
CC acids, nucleic acids, vitamins, saccharides and organic acids,
CC particularly L-lysine. The present sequence is a nucleic acid described
CC in the exemplification of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the European Patent Office
XX

SQ Sequence 759 BP; 211 A; 209 C; 189 G; 150 T; 0 U; 0 Other;

Query Match 11.3%; Score 671.8; DB 5; Length 759;

Best Local Similarity 93.7%; Pred. No. 2.4e-191;

Matches 711; Conservative 0; Mismatches 47; Indels 1; Gaps 1;

QY 1548 ATGGAATATCTATCTGCAAAAGCAGCAAGAGTGGCAAAAGCACTTGCACTGCTAATC 1607
DB 1 ATGGACATCATCATCTGCAAAAGCAGCAAGAGTGGCAAAAGCACTTGCACTGCTAATC 60
QY 1608 GCACCTTGGCAACCAAGGGTGAACCTTGGGGCTTGCAACAGATCCTCACCACTGAGT 1667
DB 61 GCACCTTGGCAACCAAGGGTGAACCTTGGGGCTTGCAACAGATCCTCACCACTGAGT 120
QY 1668 ACCTACCAAGAGCTCATTTGCAATGATGAAGCTGGGGAAGTGTCAATTCAGAACTGCAAG 1727
DB 121 ACCTACCAAGAGCTCATTTGCAATGATGAAGCTGGGGAAGTGTCAATTCAGAACTGCAAG 180
QY 1728 GCATCTTGTGGATGAATAGCTGGGACTAACCCGTCGATGAAAAAGCTACTTTAA 1787
DB 181 GCATCTTGTGGATGAATAGCTGGGACTAACCCGTCGATGAAAAAGCTACTTTAA 240
QY 1788 ACCATTGCAAGAGTCACTGACCAATGCAATCGTGTGATGAAGAGTCTACAGCCCA 1847
DB 241 ACCATTGCAAGAGTCACTGACCAATGCAATCGTGTGATGAAGAGTCTACAGCCCA 300
QY 1848 GATGCTGCAAAAGCTGATCATACGAGAGCTGCAAGATGATGAGCAAAAGATCCCTGCA 1907
DB 301 GATGCTGCAAAAGCTGATCATACGAGAGCTGCAAGATGATGAGCAAAAGATCCCTGCA 360
QY 1908 GAATCCGTTGAAGTTCAAAATCCTTGGCATCGCGGAAACGG-CACATCGCTTTCAATGAA 1966
DB 361 GAATCCGTTGAAGTTCAAAATCCTTGGCATCGCGGAAACGG-CACATCGCTTTCAATGAA 420
QY 1967 CCATCATCTCTCTGTCAGGAGTGAACAAAGTCCAGGCGCTGCACCTTAAACCTGTGAG 2026
DB 421 CCATCATCTCTCTGTCAGGAGTGAACAAAGTCCAGGCGCTGCACCTTAAACCTGTGAG 480
QY 2027 GACAAAGCTGATCTTCAACACCAATGCAAGAGTCCCAACCAAGCGCTGACCCAGGGT 2086
DB 481 GACAAAGCTGATCTTCAACACCAATGCAAGAGTCCCAACCAAGCGCTGACCCAGGGT 540
QY 2087 TTGGGCACTTTGTCCCGCGCGCAAAACATGTTGTGTGGCAACTGTGAGAAAAAGCC 2146
DB 541 TTGGGCACTTTGTCCCGCGCGCAAAACATGTTGTGTGGCAACTGTGAGAAAAAGCC 600
QY 2147 GACGCGCATCCGCGGAACCTGTGGAAGGCCAGTGAAGTCTTGGCCCAAGTTCCATCCCTG 2206
DB 601 GACGCGCATCCGCGGAACCTGTGGAAGGCCAGTGAAGTCTTGGCCCAAGTTCCATCCCTG 660
QY 2207 TAGATGCAACACATGCAATCATGTTGATGAAGCAGAGTATCCAAAGCTGGAAGAAC 2266
DB 661 CAGATGCAACACATGCAATCATGTTGATGAAGCAGAGTATCCAAAGCTGGAAGAAC 720
QY 2267 GCTGATCACTACCGTCTCATGAGCAATTAAGCTGCGC 2305
DB 721 GCTGATCACTACCGTCTCATGAGCAATTAAGCTGCGC 759

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 09:56:08 ; Search time 943 Seconds
(without alignments)
10357.302 Million cell updates/sec

Title: US-10-019-284B-1

Perfect score: 5969

Sequence: 1 agtcgcgtgcagccacacat.....tcgatatctgcgcgtctga 5969

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 81818359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3687	61.8	6911	1	US-08-311-174-4 Sequence 4, Appli
2	232.2	3.9	2913	4	US-09-710-279-3464 Sequence 3464, Ap
3	183.2	3.1	1452	4	US-09-489-039A-5847 Sequence 5847, Ap
4	181.4	3.0	1386	4	US-09-489-039A-6203 Sequence 6203, Ap
5	181	3.0	1491	4	US-09-134-000C-2417 Sequence 2417, Ap
6	177.6	3.0	30246	4	US-08-956-171E-56 Sequence 56, Appl
7	173.6	3.0	30246	4	US-08-781-986A-56 Sequence 56, Appl
8	177.6	2.9	1968	4	US-08-583-110-297 Sequence 297, Appl
9	172.2	2.9	3895	3	US-08-961-527-201 Sequence 201, Appl
10	167.8	2.8	1884	4	US-09-583-110-463 Sequence 463, Appl
11	167.8	2.8	1932	4	US-09-107-433-647 Sequence 647, Appl
12	162.2	2.7	1428	4	US-09-489-039A-1806 Sequence 1806, Ap
13	154.2	2.6	9769	3	US-08-961-527-30 Sequence 30, Appl
14	139	2.3	732	3	US-09-134-000C-1277 Sequence 1277, Ap
15	124.8	2.1	714	4	US-09-107-532A-378 Sequence 378, Appl
16	118.8	2.0	1905	4	US-09-543-681A-2847 Sequence 2847, Ap
17	118.6	2.0	1593	4	US-09-107-433-2566 Sequence 2566, Ap
18	117.8	2.0	951	4	US-09-134-000C-2077 Sequence 2077, Ap
19	117.6	2.0	1971	4	US-09-489-039A-1041 Sequence 1041, Ap
20	108.6	1.8	3615	1	US-08-920-827-17 Sequence 17, Appl
21	108.6	1.8	3615	1	US-08-921-177-17 Sequence 17, Appl
22	108.6	1.8	3615	1	US-08-921-177-17 Sequence 17, Appl
23	108.6	1.8	3615	1	US-08-921-177-17 Sequence 17, Appl
24	108.6	1.8	3615	2	US-08-920-828-17 Sequence 17, Appl
25	106.4	1.8	741	4	US-09-134-000C-2317 Sequence 2317, Ap
26	104.4	1.7	804	4	US-09-489-039A-6023 Sequence 6023, Ap
27	102.6	1.7	2230	4	US-09-949-016-4239 Sequence 4239, Ap

28	99.8	1.7	1446	4	US-09-543-681A-984 Sequence 984, Appl
29	95	1.6	465	2	US-08-673-190A-3 Sequence 3, Appli
30	92.4	1.5	1887	4	US-09-107-532A-1634 Sequence 1634, Ap
31	92	1.5	2295	3	US-09-221-017B-736 Sequence 736, Appl
32	90.6	1.5	834	3	US-09-367-293-1 Sequence 1, Appli
33	90	1.5	1830121	4	US-09-557-884-1 Sequence 1, Appli
34	90	1.5	1830121	4	US-09-643-990A-1 Sequence 1, Appli
35	88.6	1.5	708	4	US-09-583-110-829 Sequence 829, Appl
36	87	1.5	970	4	US-09-270-767-15150 Sequence 15150, A
37	84.4	1.4	840	3	US-09-543-681A-1005 Sequence 1005, Appl
38	84.4	1.4	4403765	3	US-09-103-840A-2 Sequence 2, Appli
39	84.4	1.4	4411529	3	US-09-103-840A-2 Sequence 1876, Ap
40	73.8	1.2	1284	4	US-09-107-532A-1876 Sequence 1876, Ap
41	73.2	1.2	465	4	US-08-956-171E-1317 Sequence 1317, Ap
42	73.2	1.2	465	4	US-08-781-986A-1317 Sequence 1317, Ap
43	72.2	1.2	1501	4	US-09-774-528-231 Sequence 231, Appl
44	71.2	1.2	246	4	US-09-107-433-202 Sequence 202, Appl
45	71.2	1.2	2013	4	US-09-134-000C-1201 Sequence 1201, Ap

ALIGNMENTS

RESULT 1
Sequence 4, Application US/08311174
Patent No. 5556776
GENERAL INFORMATION:
APPLICANT: TSUCHIYA, MAKOTO
APPLICANT: MIWA, KIYOSHI
TITLE OF INVENTION: SUCRASE GENE DERIVED FROM CORNYNEFORM
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT,
ADDRESSER: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/311,174
FILING DATE: 23-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 046836/1992
FILING DATE: 04-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NO. 5556776man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 10-699-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 6911 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
US-08-311-174-4
Query Match 61.8%; Score 3687; DB 1; Length 6911;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY	1	AGTCGCTGAGCGCCACCATGTGATGTGTGTGTATCCGAGCTTGTCCGAGGCTTTCTACATCT	60
Db	1	AGTCGTGTGAGCGCCACCAATGTATGTGTGTGTATCCGAGCTTGTCCGAGGCTTTCTACATCT	60
OY	61	ACGCTCCCGTCGGCGCTGTGAGTGGGGTCAATTACGGGTGGGATCAGCCCGGTGAAAGTTGG	120
Db	61	ACGCTCCCGTCGGCGCTGTGAGTGGGGTCAATTACGGGTGGGATCAGCCCGGTGAAAGTTGG	120
OY	121	GAACCCATGTGTCTCTGTGTGGGTGTGAGGAAAGAGTGTGGGTGTGAAGATTCTTCAAGTG	180
Db	121	GAACCCATGTGTCTCTGTGTGGGTGTGAGGAAAGAGTGTGGGTGTGAAGATTCTTCAAGTG	180
OY	181	TCTGTAGTTTTTAAGTTATGTATGCATCATCACTTGAGAGGTGAGGATTAATTCACTAGACCTG	240
Db	181	TCTGTAGTTTTTAAGTTATGTATGCATCATCACTTGAGAGGTGAGGATTAATTCACTAGACCTG	240
OY	241	CAACAGCAGGCTCTCAAGTCCGAGATTAATTAACTTAGATCCGTAGACATPAAGACATCAT	300
Db	241	CAACAGCAGGCTCTCAAGTCCGAGATTAATTAACTTAGATCCGTAGACATPAAGACATCAT	300
OY	301	CGTCTTAATGCTTGTCTGTGAAAGGAAACCAATTAACCTAGAAAGATGTGCAGAAAGTGTGCATT	360
Db	301	CGTCTTAATGCTTGTCTGTGAAAGGAAACCAATTAACCTAGAAAGATGTGCAGAAAGTGTGCATT	360
OY	361	ATCAAGAAATATGCSAGTCAAGCATTTAAAAAATTGAGGGAAGAAATTGTTCCCCCTCTG	420
Db	361	ATCAAGAAATATGCSAGTCAAGCATTTAAAAAATTGAGGGAAGAAATTGTTCCCCCTCTG	420
OY	421	GGGTGATTGATGTGCTTTCTCCACTGTGAAACCGGCATCATACGGAATCTCTGTGAGAAC	480
Db	421	GGGTGATTGATGTGCTTTCTCCAACTGTGAAACGGGCATCATACGGAATCTCTGTGAGAAC	480
OY	481	CAGCAGCTAAAAACGAGGATTCACCCCGAACTCCCAAGATTGTTCCCGGTTTATATG	540
Db	481	CAGCAGCTAAAAACGAGGATTCACCCCGAACTCCCAAGATTGTTCCCGGTTTATATG	540
OY	541	ATCTTCATTAATCAGGTGTGAAAACGGTGTGGCGCTTTCTTACGGGAAACGACAGCACGCGA	600
Db	541	ATCTTCATTAATCAGGTGTGAAAACGGTGTGGCGCTTTCTTACGGGAAACGACAGCACGCGA	600
OY	601	GGAACACCGCGAGATTCACCGCGAACATGTGGACGACGCTGATGTGGCAACACATGCTTT	660
Db	601	GGAACACCGCGAGATTCACCGCGAACATGTGGACGACGCTGATGTGGCAACACATGCTTT	660
OY	661	CGGCGCCGGCTGACGCACTGTGAGCGCAGGTGTGAAAACCTTATTTCCCTGTGTGAAAGAG	720
Db	661	CGGCGCCGGCTGACGCACTGTGAGCGCAGGTGTGAAAACCTTATTTCCCTGTGTGAAAGAG	720
OY	721	TCTGTCTGTGCGGCAATTCACTGTGAGGGCCCTTTTCAACAAGCATGTCCGTGTGTGCTC	780
Db	721	TCTGTCTGTGCGGCAATTCACTGTGAGGGCCCTTTTCAACAAGCATGTCCGTGTGTGCTC	780
OY	781	AAAAACCCGGATTTCAATTTTTCCCGGCAACCCAAACAGATCTTSCCGGGGATCCCAATGCGG	840
Db	781	AAAAACCCGGATTTCAATTTTTCCCGGCAACCCAAACAGATCTTSCCGGGGATCCCAATGCGG	840
OY	841	GAAGAGTTGATCAAAATCGATTCACAGTAGCGCGGAAACTGACAACTTTCTGAGCTTC	900
Db	841	GAAGAGTTGATCAAAATCGATTCACAGTAGCGCGGAAACTGACAACTTTCTGAGCTTC	900
OY	901	TTCGATCTCTGTGCGCAGCGACCAATCATATGCTTCTCTTGGGGACACTGATGTGAGATTTTG	960
Db	901	TTCGATCTCTGTGCGCAGCGACCAATCATATGCTTCTCTTGGGGACACTGATGTGAGATTTTG	960
OY	961	ATACCACTACAGCCCAATGTGCTTGGCTTAAAGAAAGAAATGTGAACGTACAGGCTACGC	1020
Db	961	ATACCACTACAGCCCAATGTGCTTGGCTTAAAGAAAGAAATGTGAACGTACAGGCTACGC	1020
OY	1021	ATTGTGTCAATGTGATGCTTCGCTGTGCATCATAGGGCTCCGGCAGCGTGTGGCGCTTTGC	1080
Db	1021	ATTGTGTCAATGTGATGCTTCGCTGTGCATCATAGGGCTTCGCCGACGCTGTGGCGCTTTGC	1080

QY	1081	TTGCTGCGGCGACGCGCCGGGGACGCATATGTTGAGTTGATGCGCGAAGGGGTGATTTGG	1140
Db	1081	TTGCTGCGGCGACGCGCCGGGGACGCATATGTTGAGTTGATGCGCGAAGGGGTGATTTGG	1140
QY	1141	CCGATGGAACGGTGCATCTAGTCTGTTCCAAACAACGCGCTTTTTCATCAGCGAGCCCATGG	1200
Db	1141	CCGATGGAACGGTGCATCTAGTCTGTTCCAAACAACGCGCTTTTTCATCAGCGAGCCCATGG	1200
QY	1201	AAGCCGCGGAAATGCCAGACGGGTGATGATCTTTTGGCGTTTGAACGTCAACGTCACCG	1260
Db	1201	AAGCCGCGGAAATGCCAGACGGGTGATGATCTTTTGGCGTTTGAACGTCAACGTCACCG	1260
QY	1261	ATGGAAGTCGCCCCGTCGCGATGCGCGGCGGCATCGCGGGGGACACGACGACTAGCGA	1320
Db	1261	ATGGAAGTCGCCCCGTCGCGATGCGCGGCGGCATCGCGGGGGACACGACGACTAGCGA	1320
QY	1321	GTCAGTTCGTCGACACAGTCGCGCAGGGGGTATGACGTATTCGACGCGACCCCTCACACT	1380
Db	1321	GTCAGTTCGTCGACACAGTCGCGCAGGGGGTATGACGTATTCGACGCGACCCCTCACACT	1380
QY	1381	CAACCGTCGCGCTAAATTTCTCGTCTTGGCGATCAGAAATGCGTAAATCCAAACCTTG	1440
Db	1381	CAACCGTCGCGCTAAATTTCTCGTCTTGGCGATCAGAAATGCGTAAATCCAAACCTTG	1440
QY	1441	CAAAATTTTGTGTCCTTTGACTCCAAACGGCCAGGTGCAAAAGTCCATTTAGTGTATCAAG	1500
Db	1441	CAAAATTTTGTGTCCTTTGACTCCAAACGGCCAGGTGCAAAAGTCCATTTAGTGTATCAAG	1500
QY	1501	TACTTTAAGTACGAGTAAACTATCTCGATTTTTPAAAGAGTCCACCATGGAATCACTA	1560
Db	1501	TACTTTAAGTACGAGTAAACTATCTCGATTTTTPAAAGAGTCCACCATGGAATCACTA	1560
QY	1561	TCTGCAAAAGACGACCAAGAACTCGGCAAGACGTTGCAGTCTTAATGCAACCTTTCGCA	1620
Db	1561	TCTGCAAAAGACGACCAAGAAAGTTCGGCAAGACGTTGCAGTCTTAATGCAACCTTTCGCA	1620
QY	1621	ACAAAGGTGGAACCTTGGGGCTTTCGACACAGATCTCTCAACATGAGTACCTTACCAAGC	1680
Db	1621	ACAAAGGTGGAACCTTGGGGCTTTCGACACAGATCTCTCAACATGAGTACCTTACCAAGC	1680
QY	1681	TCATTCGCATGTATGAAGCTGGGGGAAGTGCATTCAAGAACTGCAAGGCACTTGTGTGG	1740
Db	1681	TCATTCGCATGTATGAAGCTGGGGGAAGTGCATTCAAGAACTGCAAGGCACTTGTGTGG	1740
QY	1741	ATGAATACTGGGGACTAAACCCGTGACGATGAAAAACACTACTTTAAAAACATTCGCAAG	1800
Db	1741	ATGAATACTGGGGACTAAACCCGTGACGATGAAAAACACTACTTTAAAAACATTCGCAAG	1800
QY	1801	AGTTCACCTGACCAATGACATCGTTGATGAAAGGTCCTACAGCCCAAGATGGTGCAAAC	1860
Db	1801	AGTTCACCTGACCAATGACATCGTTGATGAAAGGTCCTACAGCCCAAGATGGTGCAAAC	1860
QY	1861	CTGATTCGATTCGAAAGCGCTGCAGAGTATGAGGCAAAAGATCGCGCGCAATTCGTTGAAG	1920
Db	1861	CTGATTCGATTCGAAAGCGCTGCAGAGTATGAGGCAAAAGATCGCGCGCAATTCGTTGAAG	1920
QY	1921	TTCAAAATCTTGGCATCGCGGAAACGGCAATCGCTTGATTTGATGAACCATCATCTTCTCT	1980
Db	1921	TTCAAAATCTTGGCATCGCGGAAACGGCAATCGCTTGATTTGATGAACCATCATCTTCTCT	1980
QY	1981	GTCAGAGCTGACAAAGGTCCTAGCGCTGCACCTCTAAACTGTGGAGGACAAACGTCGATT	2040
Db	1981	GTCAGAGCTGACAAAGGTCCTAGCGCTGCACCTCTAAACTGTGGAGGACAAACGTCGATT	2040
QY	2041	CTTCAACAACCATCGAAGAAGTCCCAACCCAGCCGTCATCCCAAGGTTTGGGCACTTGTCT	2100
Db	2041	CTTCAACAACCATCGAAGAAGTCCCAACCCAGCCGTCATCCCAAGGTTTGGGCACTTGTCT	2100
QY	2101	CCGCGCGCAAAACATCGTGTGTGTGGAGCACTGGTGAAGAAAAAGCCGACCCATCCGCGG	2160
Db	2101	CCGCGCGCAAAACATCGTGTGTGTGGAGCACTGGTGAAGAAAAAGCCGACCCATCCGCGG	2160
QY	2161	AACGTGGAAAGCCCAAGTACTGTTTTCGCCAGGTTCCATCTCTGTAGTGCACAACAT	2220

```

Db      2161  |||||
      2221  |||||
      2221  |||||
      2281  |||||
      2281  |||||
      2341  |||||
      2341  |||||
      2401  |||||
      2401  |||||
      2461  |||||
      2461  |||||
      2521  |||||
      2521  |||||
      2581  |||||
      2581  |||||
      2641  |||||
      2641  |||||
      2701  |||||
      2701  |||||
      2761  |||||
      2761  |||||
      2821  |||||
      2821  |||||
      2881  |||||
      2881  |||||
      2941  |||||
      2941  |||||
      2941  |||||
      2941  |||||
      3001  |||||
      3001  |||||
      3061  |||||
      3061  |||||
      3121  |||||
      3121  |||||
      3181  |||||
      3181  |||||
      3241  |||||
      3241  |||||

```

```

Db      3241  |||||
      3301  |||||
      3301  |||||
      3361  |||||
      3361  |||||
      3421  |||||
      3421  |||||
      3481  |||||
      3481  |||||
      3541  |||||
      3541  |||||
      3601  |||||
      3601  |||||
      3661  |||||
      3661  |||||

```

RESULT 2
 US-09-710-279-3464/c
 ; Sequence 3464, Application US/09710279
 ; Patent No. 6703492
 ; ORGANISM: KIMBERLY, WILLIAM JOHN
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 ; FILE REFERENCE: PUS480US
 ; CURRENT FILING DATE: 2000-11-09
 ; PRIOR APPLICATION NUMBER: 60/164,258
 ; PRIORITY FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 4472
 ; SOFTWARE: Patencin Ver. 2.1
 ; LENGTH: 2913
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic
 ; OTHER INFORMATION: nucleic acid sequence
 US-09-710-279-3464

Query Match 3 9%; Score 232.2; DB 4; Length 2913;
 Best Local Similarity 54.0%; Pred. No. 5.8e-62;
 Matches 574; Conservative 0; Mismatches 468; Indels 21; Gaps 4;

```

      4053  |||||
      2903  |||||
      4113  |||||
      2843  |||||
      4173  |||||
      2783  |||||
      4230  |||||
      2723  |||||

```

```

QY 4290 CAGTTGGTGGTTCACCGCAACGAGTTCCGTGGCAATGAGTTCCTGGGCGCG 4349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2663 CAATACCTTATTTGGTTTATGAGCAGCTAAGCAGATTGGTGGTATCCTTATTTAGGTGAG 2604
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4350 GCAATGGTATGGGATGGTGGTCCCAACCTGGTTAAGGCTAAGCAGCGGGCGCCACA 4409
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2603 CTTAGAGTATGATCTTGTTCATCTCGAATGATGATGATGATGATTTTCCAAAAGCTT 2544
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4410 TGAACCGGCGGCGAA--ATGCCAATGGTCCCTGTTGGTGGATGGTCTCAAGCTG 4466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2543 TAGAAGAAAGAAAGCTATTTCCACACTGGGATGCTTTGGGCTACATATTATGAAGTGA 2484
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4467 GTTACAGGAGCAGCCGCTCTCTGTGCTGTGTCTCTTGATTTGGCAAGATCGAGA 4526
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2483 GTTATCAGGAGCAAGATTAATCTATGCTGTGTGCAACATATATTTAGCTACGATTTGAA 2424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4527 AGTTCCTGACAAAGCACTCATGGGCACTGCAGACTTCTGATCAACCCAGTGGTGAATC 4586
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2423 AATGCTTACGTAAAGTATTTCCACCTGTGTATGATATTTATTTAGCCCATTTATTCAA 2364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4587 TCTGCTACCGGCTTCTTACGTTCAATGCTATTGGTCACAAATGCGCTGGGTGGT 4646
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2363 TTTTATTTACAGCATTTATTAACATTTTATTTGAGGCGCTGTCACTGTCATTAGGTT 2304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4647 ACTTGCTGACAGCGTCTGCAGGAGCTGATGATTTGGTGTCCAGTGGCGGTGTC 4706
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2303 ATTTGTTATCTGATGATTTGATGCTGTATATGAAATTTGTGAGCTATTTGATGACTTA 2244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4707 TTTTGGTCTGATCTACTACCAATCGTATGCTATGCTGTGACAGCAAGTCTTCCGCCAA 4766
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2243 TATTTGGTTATTAATATGCGCAATGCTATTAAGGAATGACACATAGCTTTATTTGAG 2184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4767 TTGAGCTG-----GAGCTTTCAACAGGTTGATCTTCATCTTTCGACACG 4814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2183 TTGAAACGACATTAATGCTGATGCGACATMAAACGTTGTCATTTATCTTCCCAATCG 2124
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4815 CATTCATGCGCAATATCGCGAGGGTGCAGAGTGTGGCAGTGTCT---TCCTAGGA 4871
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2123 CAACATGTCAAATTTGACAAAGGTGTGACAGCTTTAGCTCAATCTTTATCATTTAAGC 2064
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4872 AGAGTGAAGAGCTCAAGGCGCTTGACAGTGTCTTCAGTGTCTCCGCTGTTCTTGTGATTA 4931
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2063 AAAATATAAAATTTAAAGGTGTGTGCTCGCGCGGGAATTTCACTTTACTAGGAATTA 2004
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4932 CAGAGCTGCGATCTTCGTGTGAACCTTGCCTGCGCTGCGCTTCAATGTGATCG 4991
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2003 CAGAACCAAGCAATGTTGTGTCAATCTTAAATGAATATCATTTAAGGTGCTGTG 1944
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4992 GTACGCAAGCTATCGGTGCGCTTGTGATGCACTTTGATATCAAGGAGTGTGGTGG 5051
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1943 CAGGATCAGGATAGGTGGGCTTATATTCTTCAAGTAAAGGATAGGCGCTTG 1884
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 5052 GCGCTGCAAGTTCCTTGGGTGTGTTCTATTGATGCTCCAGA 5094
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1883 GTACAGCTGATTAACCTGATTTATATCTATAATCTACACA 1841
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 3
US-09-489-039A-5847
; Sequence 5847, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709, 2004001
; CURRENT APPLICATION NUMBER: US/09/489, 039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5847

```

```

; LENGTH: 1452
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-5847

```

```

Query Match 3.1%; Score 183.2; DB 4; Length 1452;
Best Local Similarity 48.4%; Pred. No. 1.2e-46;
Matches 678; Conservative 0; Mismatches 668; Indels 54; Gaps 4;

```

```

QY 3754 TTTCTCCCGTTAAGAGAAATTCATGACCATATAGACCTCGGCAACCATCTCG 3813
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 57 TATATCATGATTTAGAGGATCAATGATTTTGAAGATTTCCGCTCACTGCTTCC 116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3814 CGACATTTGGCGGAGAAACAATTTGCGCCGCGCACACTGTGCAACGCTTTAGCCT 3873
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 117 CTTGCTGGCGGCAAGAAATATCGCAGGCGCGCGCATGCGCACCCGCTCGGCT 176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3874 CGTGTCAAAAGACCAAGATGTGATCGCCAAAGTGTGATGATTCAGATCTGAA 3933
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 177 GGTGTGATGACAGCGGCTGCGGATGACAGGCGATTTGCAAAATGACGGGGTGA 236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3934 AGGCACTTTGAACTGCGGCAATGTTCCAGATCATCTCGGCGCAGGCGATGTGATCA 3993
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 227 AGGCTGCTTTGCAATGCGGACAGATGACATATCTTTGGGACCGGGGTGTCAATTA 296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 3994 TGTTTCAAAAGACTGATGACGCACTCCAAAGACTGCTGTGTTCAGAGACGCT 4053
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 297 AG--TCTATGCCGCTTTATCAAGCGCGAGGATCAGCAATCGAGCAATCGAAGC 353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4054 CAAGAATTTGTGCTAACAAGCAAGCACTGTGAGCGGCTGTGAAAGTATTTGGCGA 4113
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 354 CGCGACCTGCGGGGAAAGCTGAAAGCCGTTTCAGAGCATGCCCCGCTGTGTCCA 413
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4114 CATTTGCTCCGCTGATTTCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4173
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 414 CATCTTCGTGCGATTTATTTCCGCGCATGCTGTGCTCGGCGCTGTGATGGCGCTGG 473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4174 TGTGTGTGTGCGAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 474 GATGTGAAACCTTAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 506
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4234 CAGCGTGTGTGCTAGATGATCACTGATGAGCATCTGCGCGCTTCTTGGCCAGT 4293
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 507 CAAGCTCTCATATCATGCTGATATGTGATGTGAGTTCGGCGGCTTTATCATCTGCGAT 566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4294 GTTGTGTGTGTCAACCGCAACAAAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4353
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 567 CTTATCGGCTTTTACCGCGCAGCGAATTTGGCGGTAAACCTTATCTGGCGGACCT 626
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4354 TGTATGCGCATGATGTTCCCAACCTGTGTTAAGGCTAAGAGTGGCGCGCACATGAC 4413
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 627 CCGGCGGATCTCAACCCATCGGCGCTGACCAAGCGCTGGGCGGTGCGCG----- 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4414 CGCGGCGAAATGCGCAATGTGTCTCTGTTGTGTGTGTGTGTGTGTGTGTGTGTGT 4473
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 678 -----CGGCTTCAACCAATGATTTCTTGCGATGAGTGTGCAATGCGCTACA 731
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4474 GCGCACTGTGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4533
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 732 GGGGACCGTCTTCCCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4534 GCACAAGGACTCAATGGGCACTGCAGACTTCTGATCAACCCCAAGTGTGACTGTGCT 4593
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 792 GCGCGCGTATCTCTTACGCGCTGACGCTGATCTTCACTTCTGCTGAGCGATTTAT 851
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4594 CACGCGCTTCTTACGTTCAATGCTATGTGTGTCAGCAATGCGCTGTGTGTGACTTGTCT 4653
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 852 CTCGCGCTTATGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 911
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4554 GGCACAGGCTGTGAGGAGATCTATGATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4713
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 912 TTGCTTATCTTCAGACGCTTATCAGCCAGCGCGGTGCTGTGCGGCTGTGCTTTCG 971
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


COMPUTER READABLE FORM:

```

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: P8248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 30246 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 56:
US-08-956-171E-56

```

Query Match	3.0%;	Score 177.6;	DB 4;	Length 30246;
Best Local Similarity	49.1%;	Pred. No. 9.5e-44;		
Matches 607;	Conservative 0;	Mismatches 599;	Indels 30;	Gaps 4;
Qy	3801	AACGCATCTGCGCGACATTGGCGGCCAAGACA	CATTTGTGCGCGCGCACACTGTGGCA	3860
Db	10819	AAGCCATGTATAACCGCATTTGGGGGAAAAA	CTTTGAAGGTGCAACGCACTTGTGTAA	10876
Qy	3861	CGCGTTTACGCGTGTGTCTCAAAGACCA	AGATGTGGATTCGCCAAAGTCTGGATGATG	3920
Db	10879	CACATTACGTTTATGTCTGAAGGATGAAGTAA	AGTTGATTAAGACGCAATTAAAGTAATA	10938
Qy	3921	ATTCAGATCTGAAGAGCACTTTGAAATCTG	ACCGGCAATGTTCCAGATCATGTGCGGCGAG	3980
Db	10939	ACGCGTTTGGTCAGAGGGGCACTTTAAAC	AGACCACTCATATTCAAATTTGTATTGTGTCA	10999
Qy	3981	GCGATGTGATCATGTTTCAAAGAACTCGAT	AGCAAGCACTCCAAAGACATCGCTGTGT	4040
Db	10999	GAAACATCGATGAAGTGTATTAAGACGTTT	ATTGATGAACAGGNGCTCAAGAAAGCT---	11055
Qy	4041	CCACAGACAGCTCAAGATGTTGTGGCTAA	CAACCGCACTGTTCAGCGGCTGTGTGA	4100
Db	11056	CGAAAGATGAAGCGAAACAAAGCACTGCA	CAAAAAGGAAATCAAGTCAACGTTTGATCA	11115
Qy	4101	AGGATTTGGCGGACATTTTGGTCCCGCTGA	ATTCATCTTGTTGGTGGCGGCTGTGCTCA	4160
Db	11116	AATTGTTGGGGGAAATTTTATACATATTA	CTCGGATTTGTGAACAGCGTGTGTGTAA	11175
Qy	4161	TGCGTATCAACAATGTGTGTGGTTCGCGAG	ATCTGTTCCGCAATCATCTGTGGAGA	4220
Db	11176	TGGGAATCAATTAATTTACTTACAATGAAG	GTTTATTTGGTCCAAAGCACTTAATTGAGA	11235
Qy	4221	TGTTCCCTCAGATAGAGGATGTGGTCTAGA	ATGATCAACTGATGCGCATCTGGCGCGGTG	4280
Db	11236	TGTAATCAACAATTTGCTGATATTTCAAAC	ATCATTAATGTGAATTTGCGATGCGGCA	11299
Qy	4281	CGTTCTTGCCAGTGTGGTGTGTTTCAAC	CGCAACCAAGCGTTTCGTTGCGGCAATGA	4340
Db	11296	TTTTCCTTACAGCATTAATTTGGTTGAG	ATGATGCGTGTATTTGGGGTGAATGCCATTC	11355
Qy	4341	TGGGCGCGCGCATTTGATGCGCATGCTAT	TCCCAACCTCGGTTAAACGCTACAGACGTG	4400
Db	11356	TAGGCAATGCTTAATGTTTAAATGATCCGA	CAATTAATGATCTCAGTAAATGATTTGG	11415

RESULT 7

```

US-08-781-986A-56
: Sequence 56, Application US/08781986A
: Patent No. 6737248
:
: GENERAL INFORMATION:
: APPLICANT: Charles Kunsch
: TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
: NUMBER OF SEQUENCES: 5255
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII text
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/781,986A
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
:

```

```

1  ATTORNEY/AGENT INFORMATION:
2  NAME: Benson, Bobb
3  REGISTRATION NUMBER: 30,446
4  REFERENCE/DOCKET NUMBER: PB248PP
5  TELECOMMUNICATION INFORMATION:
6  TELEPHONE: (301) 309-8504
7  TELEFAX: (301) 309-8512
8  INFORMATION FOR SEQ ID NO: 56:
9  SEQUENCE CHARACTERISTICS:
10  LENGTH: 30246 base pairs
11  TYPE: nucleic acid
12  STRANDEDNESS: double
13  TOPOLOGY: linear
14
15  JS-08-781-986A-56

```

Query Match	3.0%	Score 177.6	DB 4	Length 30246
Best Local Similarity	49.1%	Pred. NO. 9.5E-44		
Matches 607	Conservative 0	Mismatches 599	Indels 30	Gaps 4

OY	3801	AACGATCTGCGCGGACATTGGCCGGCCAAACAACATTTGTCGGCCGCGACACTGTCGAA	3860
Db	10819	AAGCCATCGTAACCGCTATTGGGGGAAAABAAAATCTTGAAGCTGCAACGCAATTGTGTA	10878
OY	3861	CGCGTTTACGCTCTGTGCTCAAAAGACCAAGATGTGATCGCCAAAGTTGATGATG	3920
Db	10879	CACCATTAACGTTTATAGTGTCTAAGAGATGAAGTAAAGTTGATTAAGACGCTTAAGTAATA	10938
OY	3921	ATCCAGATCTGAAGAGCAACCTTTGAAATCTGGCCGCAATGTTCCAGATCTACGTGCGGCCAG	3980
Db	10939	ACGCGTTGGTCAAGGGGCAAGTTTAAAGCAACATCAATATCAAAATGTCATTGGTCCAG	10998
OY	3981	GCGATGTGATTCATGTTTCAAAAGAACTCGATACGCAACCTCCAAAGACATCGCTGTGT	4040
Db	10999	GAAACGTCGATGAAGTGTATTAAGCAAGTTTATGTATGAAGAACGGTGTCCAAAGAGCT--T	11055
OY	4041	CCACAGACAGCTCAAAAGATGTGTGTGCTTACAAACGCCAACTGTTCAAGCTGTCTGTA	4100
Db	11056	CGAAAGATGAAGCGAAACAAGCAGCTGCAAAAAGGGAATCCAGTACAACGTTTATCA	11115
OY	4101	AGGATTTGGGGGACATTTTGTGTCGCCGTGATTTCCATCTTTGTGTGGTGGCCGTCTCTCA	4160
Db	11116	AATGTGTGGGGGATATTTTATACCAATATTACCTGCGATTTGACAGCTGTGTGTGTTAA	11175
OY	4161	TGCGTATCAACAATGTGTGTGTGGCGAGATCTGTTCCGTCGCGAATCACTGTGTGGAGA	4220
Db	11176	TGGGAATCAATATTATCTTCAATGAAGGTTTATTTGTGTCAAAAGCACTTATTGAGA	11235
OY	4221	TGTTCCCTCAGATCAGCGGTGTGTCTGAGATGATCAACCTGATGGCATCTGGCCCGTTGG	4280
Db	11236	TGATTCACAATAATGTCTGATATTTTCAAAACATCATTTATGATGTGGAGTACGGCACTTTA	11295
OY	4281	CGTTCTTGCCAGTGTGTGTGTGTTTCAACGCCAACAAAGCTTTCCGTGTGCATGAGTTCC	4340
Db	11296	TTTCTTACACAGCATTAATGTGTGGAGTATGATGTGCTATTTGGTGTGAGTCCGATTC	11355
OY	4341	TGGCGCGCGGCAATTGATATGGCAATGCTGTTCCCAACCTCGTGTAAACGCTACGACGTGG	4400
Db	11356	TAGGCATATGCTTATAGTTTATTTAATGATCCGCAATTAAGTATCTAGTATGATTTGG	11415
OY	4401	CCGCAACCATGACCGCGGCGGAATGCAATGTGTGCCCTGTTGTGTGTGATGTTGTC	4460
Db	11416	-----CAAAAGGGAATATTCGACACGTGGAACTTATTTGGCTTAGAGATTAAAC	11463
OY	4461	AAGCTGTTTACAGAGGCAACGCTGTTCTGTGCTGTGTTCTCTTGGATTTGGCAACGA	4520
Db	11464	AGTTGAATTACCAAGGTCAAGTGTGTGCAAGTTTATTAATGACCTTAAGTTTACGTAAATA	11523
OY	4521	TCCGAAGATCTCTGACACAGCAGCTATGGGCACTGCAAGACTTCTGATCAACCCAGTGT	4580
Db	11524	TTGAAAAAGATTAAATTAAGTGTGTCAACGATTCGATTAATAAAAGTTGTGTGTGACCCG	11583
OY	4581	TGACTGTGCTGCTACCGGCTTCTCTTAAGTTCAATTTGGTTCAGAGCAATCGCGTGGG	4640

Db 11584 TAGGCGCTTTTAACTACTGGAATTTTAAAGCATTTTATATACATTTGGACGAGTTGGCTTATGA 11643

Qy 4641 TGGGTGACTTCTGTGGACACGGTCTGGACGAGACTGTATGATTTTCGATGCTCAATCGCGG 4700

Db 11644 TTGGTACAGGTATTAATCTGTGGTGTTCATTTATATATTCGAATGACGAGATGGCTTGGCG 11703

Qy 4701 GTCTGCTTTTGGGTCTGGTCTACTGACCAATGTTATATCATCTGTGGACCAAGCTCTTCC 4760

Db 11704 GAGCAATATATGATTTGTTATATGACCACTTGTAATTAACAGACACTACACCAATATGTTTT 11763

Qy 4761 CGCCAAATGAGCTGGAGCTGTT-----CAACGAGGTGATTCCTTCACTTCGCAACCG 4814

Db 11764 TAGCAGTATGATTTCCAAATTTGATGGAGTACAGCTTACGGCTATGATTTATATGCGCAATTG 11823

Qy 4815 CATCATGGCCCAATATCCGCGAGGGTGCAGCATGTTTGGCAGTGTCTTCTACGCAAGA 4874

Db 11824 TTGCATTTTCCAAATTTTGTCTGAGGCTCTGCAGACATTTGGACATGTTTGTCTATAAAC 11883

Qy 4875 GT-----GAAAGCTCAAGGGCTTGCAGGTGCTTCAAGTGTCTCCGCTGTCTTG 4925

Db 11884 GTCTTAATAATGTTTAAAGAGAGAGGCTTGGCAATTAACATCTTGATTTCTGGTATGTTAG 11943

Qy 4926 GTATTACAGAGCTCGATCTTCGGTGTGAACTTCCTGCTGCGCTGGCGGCTTCTAATG 4985

Db 11944 GTGTTACTGAAACACCGCAATGTTGGGTGATTAATCTCTGAAATATCATTTATTCGCTG 12003

Qy 4986 GTATCGTACCGCAGCTATTCGGTGGCGCTTTGATTG 5021

Db 12004 CGATATCAACGCTTGTGTATTTGGGGCAATGCTTG 12039

RESULT 8

US-09-583-110-297

/ Sequence 297, Application US/09583110

/ Patent No. 669703

/ GENERAL INFORMATION:

/ APPLICANT: Lynn Doucette-Stamm et al.

/ TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus

/ TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics

/ FILE REFERENCE: PATH00-07A

/ CURRENT APPLICATION NUMBER: US/09/583,110

/ CURRENT FILING DATE: 2000-05-26

/ PRIOR APPLICATION NUMBER: US 09/107,433

/ PRIOR FILING DATE: 1998-06-30

/ PRIOR APPLICATION NUMBER: US 60/085,131

/ PRIOR FILING DATE: 1998-05-12

/ PRIOR APPLICATION NUMBER: US 60/051,553

/ PRIOR FILING DATE: 1997-07-02

/ NUMBER OF SEQ ID NOS: 5322

/ SEQ ID NO 297

/ LENGTH: 1968

/ TYPE: DNA

/ ORGANISM: Streptococcus pneumoniae

US-09-583-110-297

Query Match 2.9%; Score 173.8; DB 4; Length 1968;

Best Local Similarity 46.5%; Pred. No. 1.6e-43;

Matches 876; Conservative 0; Mismatches 962; Indels 45; Gaps 8;

Qy 3797 GCGCAAGCATCTCGCGGAGACATTGGCGGAGCAACATTTGTGCGCGCGACACATGT 3856

Db 22 GCGCAAGATCTGCTTCAGGCAATGGAGCAAGAAATGTGACTGCGGTAACTCACTGT 81

Qy 3857 GCAACGCGTTTAAAGCGCTGCTCAAAAGACCAAGAGATGAGATGCGCAAGTCTGAT 3916

Db 82 GCGACACGGAATCGCTGTTGTTTAAAGATGATTAAGAGGCTAAATGTTAAAGCTATCAG 141

Qy 3917 GATGATTCAGATCTGAAAGGACCTTTGAAACTGGCGGCATGTTCCAGATCATCTGCGGG 3976

Db 142 TCAATTCAGCTGTTAAAGAACTTTTCAATGACAGCTCAATTCAGATATCAT--- 197

Qy 3977 CCAGCGCATGTGATCATGTTTCAAAGAACTCATGACGCACTCCAAAGACATCGCT 4036

	Query Match	Similarity	2.9%	Score 173.8	DB 4	Length 1968	
	Best Local	Similarity	46.5%	Pred. No. 1.6e-43			
	Matches	876	Conservative	0	Mismatches 962	Indels 45	Gaps 8
QY	3797	GGCGACGCAATCTCCGGGAATTGGCCGGAGACAACATTTGTGCCGCCGCACACTCT					3856
Dd	22	GCCAAAGATCTGCTTCAGGCAATTGGAGGCAAAAGAAAATGTACTCCGTAATCTA					81
QY	3857	GCAAGCGCTTAA CGCCTCGTGCTCAAGACA CCAAGA GTGGAT GCGCAAG TCGAT					3916
Dd	82	GCGCACGGAATCGCTTTGTTTTAAGATGATAA GAAGGCTAA TGTTAAAGCTATCGAG					141
QY	3917	GATGATCAGATCTCA AAGGCA CCGTTTGAA CTGGCGCATGTTCCAGATCA TCGTCGG					3976
Dd	142	TCAATTCAGCTGTAAAGAACTTTTACAAATGAGATCAATTTCAAGTATCAT----					197
QY	3977	CGAGCGATGTGATCATGTTTCA AAGACTCGATGACGCA CCTCCAAGACATCGCT					4036

Db 198 --TGAAATAGAGTGGCCATCTTTTAAATGATTACAGCCGTTTCAAGTAATGAGGT 255
 Qy 4037 GTGTCCAGAGAGAGTCAAAAGATGTTGGCTTAACAAGCCACTGGTTACCCCTGCT 4096
 Db 256 GTTTCAGAAAGAACCAAGTCTGCAAGTAAGATTAACAAACGTTGTTCAAGGTGTT 315
 Qy 4097 GTGAAGATATGCGGACATATTTCCGCGCTGATTCATCTGGTTGGCGGTG 4156
 Db 316 ATGACACACTCTGGCGGAATCTTTACTCCGATTAATTCAGCCCTGATTTGTTGGTGAATG 375
 Qy 4157 CTCATGCTATCAACATGTTGTTGGCGCAGAT-----CTGTTCCGTCGCAATCA 4210
 Db 316 ATCTTCGTTCCGTAATGTTGGAAGGTGTCATTTGATGATGATGATGCAAGACC 435
 Qy 4211 CTGGTGAATGTTCCCTCAGATCAGCGGTGTTGCTGAATGATCACTGATGATCT 4270
 Db 436 ATCAAGAAATCCCTCAGTTTGGGAGGGGTTAACTCTCTCGTTGCTCGTGTGA 495
 Qy 4271 GCGCGTTGCGCTTTCTTCCAGTGTGTTGGTTTACCCGCAACAAAGCTTTCCGTGAC 4330
 Db 496 GCTATCTTCAGTTCTTACCAAGTAGGATTAATCTGTTCTGTTCTGTAAGATGGAAAC 555
 Qy 4331 AATGATTCCTGCGCGCCGCAATGCTATGCGAGATGTTCCCAACCTGTTTACGCG 4390
 Db 556 ACCCAAAATTTGGGAATGTTCTCGGAATCTGTTGTAATGCTCAGTTGCTCAATGCC 615
 Qy 4391 TACGATGCGCCGCAACCATGACCGC-----GGGGAATGCAATGTTGTTCCCTGTT 4444
 Db 616 TATGGGTTGCTTCAACGCGCAGCAGTATGCAAGCAAACTGGGTTTGAATTTTGGT 675
 Qy 4445 GGTATGATGTTGCTCAAGCTGTTTACAGGGGCAACGCTGTTCTGTTGCTGTTGCTCT 4504
 Db 676 TATTTTACCTGTTATGTTATGTTATGTTTACAGCCCAAGTTATCCAGCCCTTCTGCAAGT 735
 Qy 4505 TGGATTTGGAACCAATGCAAGATTTCTGCAACAAGCACTCATGGAACCTGCAAGCTTTC 4564
 Db 736 TTGAGCTGTCTTATCTGAAATTTTGGCGCAACATATCCAGAAATGCAATTTCCATG 795
 Qy 4565 CTGATCACCCCAAGTGTGATCTGCTGCTCAACCGCTTCTTACCTTCACTTATGTTGTT 4624
 Db 796 ATTTTGAATCTTTTGTGATGATTCAGCCCTGATTTGACCTCATCTGCTTGGGG 855
 Qy 4625 CCAGCAATGCGTGGTGGTGAATCTGTCGCAACCGT---CTGAGGAGCTTATGAT 4681
 Db 856 CCAATCGTTGGAACATTTGGAAGACCTTTTCACTGTTATGTCAGAGTTTAAAGGTT 915
 Qy 4682 TTCGATGTCAGTGGCGGTCTGCTTTTCCGTTGCTGATCTCAACCAATGTTTACT 4741
 Db 916 CCTGTAATAGGCTCTTGGTCAATTTTGGTGGCCCTTACGCTCAATTTGTCATCA 975
 Qy 4742 GGTCTGCAACGATCTTCCCGCAATTTGAGCTGAGCTGTT---CAACAGGCTGATTC 4798
 Db 976 GGTCTGCAACCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 1035
 Qy 4799 TTCATCTTGGCAACCGCATTCATGCGCAATTCGCGAGGAGGATGTTTGGACAG 4858
 Db 1036 GCGCTTGGCTTATGATGCTCTTTTAAATATGCTCAAGGCTCAGCGCTGTTGCTAT 1095
 Qy 4859 TTTCTTCTAGCAAGAGTGAAGTCAAGGCGC---TTGAGAGTCTTCAAGTGTCTCC 4915
 Db 1096 TATTTTCATGATCGCATATGATGATGATGATGATGATGATGATGATGATGATGAT 1155
 Qy 4916 GGTGTTCTTGTATTAACAAGCTGCGATCTTGGTGAACCTTGGCTGGCGCGG 4975
 Db 1156 GCTATCTCGGTTTACAAACAGCTCTTTTGGGGTTAAGTAAATATATTTATCA 1215
 Qy 4976 TTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5035
 Db 1216 TTTGTTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1275
 Qy 5036 AAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5095
 Db 1276 ACTGCGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1335

Qy 5096 ATGTCATGTTCTTGGTTTGGCGGATGTTTACCTTGTGATGCAATGCGGAGGAT 5155
 Db 1336 ATGTCGCAATTTGCAAGAACTATGCTAGTTGATGTTGTTGCTCAATGCTC-----TTG 1389
 Qy 5156 GCTTATGAGCTTCTTATGTTGTTGCGCAACGCGACATTTGATGATGATGATGATGATGAT 5215
 Db 1390 ACTTCTTCTTCTGCAAGCTGCTCTTTCACAAATATGAGGGGATGATGATGATGATGAT 1449
 Qy 5216 CCAAGTCTGCAAGAACGCAAGCGAGCAAGCAACCCGCAAGATTTTCAACGAT 5275
 Db 1450 GCAGATTTGTTGCTCAAGAAAGACAGATTTGATGAGCCATGAAACAGTAACTTACT 1509
 Qy 5276 TCCACATCATCAGACGACCTTTGACCGGTGAAGCTATGCACTGACAGCGCTCAGCAT 5335
 Db 1510 TCGTAAGAAATTAATAGCCCACTAACTGCGCAAGTGAAGAAATGATGATGATGATGAT 1569
 Qy 5336 GCCATGTTTCCAGCGGAAAGCTGCTCAGGTGTTGCGATGCTGCCCAAGGAGGAG 5395
 Db 1570 CCTGTTTTCATCAGGTGCTATGAGGCAAGCTGCTGATGATGATGATGATGATGATGAT 1629
 Qy 5396 CTGTTTCAACAGTGAAGGAAAGATGCTGTTGCGCTTCCATCTGTTGATGATGATGAT 5455
 Db 1630 TTGACCTCTCAGTTAATGAAACAGTGAAGGCTTCTTTCCCAAGAGATGCAATCGAG 1689
 Qy 5456 GTCCGCACTAAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 5515
 Db 1690 ATGTTCTGACGAGG-----GAGTTGAATGCTCATCACTGCTATGAT 1737
 Qy 5516 ACCGTAACCTCAAGCGCAGCAGCTTAAACCGCTGAAGAAACAGGAGGATGATGAT 5575
 Db 1738 ACAGTATGCTTGAATGAGCAAGTTTGAAGATCTTGAATGATGATGATGATGATGATGAT 1797
 Qy 5576 GCAGGAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5635
 Db 1798 GTTGGCAGCAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1857
 Qy 5636 ACCAGCGGATGTTGTTGGA 5658
 Db 1858 GAACTCTGTTATCATCA 1880

RESULT 9
 US-08-961-527-201/C
 ; Sequence 201, Application US/08961527
 ; Patent No. 6420135
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunsch
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 391
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/961,527
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Brookes, A. Anders
 ; REGISTRATION NUMBER: 36,373
 ; REFERENCE/DOCKET NUMBER: PB340P1

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 201:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3895 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-961-527-201

Query Match 2.9%; Score 172.2; DB 3; Length 3895;
 Best Local Similarity 46.5%; Pred. No. 8.9e-43;
 Matches 875; Conservative 0; Mismatches 963; Indels 45; Gaps 8;

```

QY 3797 GCGCAACGATCTGCGGCAATGCGCGGAAACAACATGTCGCCGCGCACACTGT 3856
DB 3874 GCCAAGATCTGCTTCAAGCAATCGAGGCAAGAAATGTGATCGCTAATCACTGAT 3815
QY 3857 GCAAGCGTTACGCTCGTCTGCTCAAGAACAAGATGAGATCGCAAGTCTGAT 3916
DB 3814 GCGACACGATGCGGTTGTTTGAAGATGAAGAGCTAATGTTAAAGCTATCGAG 3755
QY 3917 GATGATCCAGATCTGAAAGGACCTTTGAACTGCGCGGATGTCAGATCATCGTCGG 3976
DB 3754 TCAATTCAGCTGTAAAGAACCTTTACAAATGACAGTCATTTCAAGTATCAT---- 3699
QY 3977 CCAGCGATGATGATCATGTTTCAAGAACTCGATGACGCACTCCAAAGCATGCT 4036
DB 3698 --TGAAATGACGTGCCCATCTTTATATGATTTTACAGCGGTTACAGTATGAGGT 3641
QY 4037 GTGTCACAGACACTCAAGATGTTGCTACAAACGCACTGTTGACCGTCT 4096
DB 3640 GTTTCAAAGAGACAGCCAGTCTGAGTAAAGTAACTAAAGTGTCCAGAGTGT 3581
QY 4097 GTGAAGTATGCGGCAATTTTCGTCGCTGATTTCCATTTGTTGTTGCGGCTCTG 4156
DB 3580 ATGACACTCTGCGGAGATTTTACTCGATTAATTCAGCTTATATGTCGAGATTTG 3521
QY 4157 CTCATGCTATCAAAATGTGTGTGCGCAGAT-----CTGTTGCGTCCGAATCA 4210
DB 3520 ATCTCGGTTCCGTAATGTCCTGGAAGATGTCATGTCATGTCATGTCATGTCAG 3461
QY 4211 CTGCGGAGATGTCCTCAGATCAAGCGGTGTCGAGATGATCACTGATGTCATCT 4270
DB 3460 ATCAAGATCTCTCAGTTTGGGAGGTGCAATCTCTCTGTTGCTGATGAA 3401
QY 4271 GCGCGCTTCGCTTCGTCAGTGTGTGTTTCAACGCAACCAAGCGTTTCGATGAC 4330
DB 3400 GCTATCTTCAAGTTCTTACAGTAGGATTAATCTGCTGTTCTCTGTAAGATGGAAC 3341
QY 4331 AATGATCTCTGCGGCGCGGCAATGATGCGATGCTGTTCCAACTGCTTAAAGCG 4390
DB 3340 ACCCAATTTTGGGAATGTTCTCGAATCTGTTGTTATCGCTCAGTGTCTCATGTC 3281
QY 4391 TACGAGTGGCGCGCACCAATGACCGGCGGCAAAATGCCAA-----TGTGTCCTGTT 4444
DB 3280 TATGCGGTTGCTTCAAGCGCAGCTATATCGCGCAAACTGGGTTTGAATTTGAC 3221
QY 4445 GGTTCGATGTTGCTCAAGCTGTTTACAGGCGACCGTCTCTCTGTCGTGTCCT 4504
DB 3220 TATTTTACGTTAATCGTATGCGTTTACCAAGCCCAAGTATCCAGCGCTGCTTCAAG 3161
QY 4505 TGGATTCGCAACGATGCAAGATTTCTGCAACAGCGACTCAATGGGCACTGCAAGTTC 4564
DB 3160 TTGAGCTGCTTATCTTGAATCTTGGCAAGCATATCCAGATCTCAATTTCTATG 3101
QY 4565 CTGATCAACCCAGTGTGATCTGCTGCTCAACGCGCTTCTTACGTTATGTCATGCT 4624
DB 3100 ATTTTGTACCTTCTGCTGATGATTCAGCGTTGATTTGCTCATATCTTTTGGGA 3041
QY 4625 CCAGCAATGCGCTGGGTGGGTACTTGTGCAACAGGT---CTCAGGACTCTATGAT 4681

```

```

DB 3040 CCAATCGGTGACATTTGACAAAGACTTTCATCAGTGTGCTTGGCAGTTTAACTG 2981
QY 4682 TTGCGGTGTCAGTTCGCGGCTCTGTTTGGTCTGCTCTACCAATGTTATCT 4741
DB 2980 CCAGTAAATGCTCTTCGGGCAATTTTGGCCCTTCAGGCTCCATTTGTATGACCA 2921
QY 4742 GATCTGACAGTCCCTTCCGCCAATTTGAGCTGAGCTGT---CAACGAGGTGATTC 4798
DB 2920 GATCTGACCAATGACCAATGCAATGATATACAAATTTGAGGAGTCTGCTGACCT 2861
QY 4799 TTCACTTCGCAACCGCATTCATGCGCAATATGCGAGGATGACGATGTTGGCAGTG 4858
DB 2860 GCCCATGACCAATGATGCTCTTTCTATATTTGCTCAAGGCTGACCGGTTTGCCTAT 2801
QY 4859 TTCTTCCAGGAAGAGTGAAGAACTCAAGGGCC---TTGAGAGTCTTCAGTGTCC 4915
DB 2800 TATTTACATGATCGCCATGATGAGCGTGAAGCTCAAGTTTACTCTGCAACCATTTCA 2741
QY 4916 GCTGTTCTTGGTATTTACAGAGCTGCGATCTTCGAGTGAACCTTCGCTGCGTGGCG 4975
DB 2740 GCTATCTCGTGTATACAGAACAGCTCTTTTGGGTTAAGTAAATATATTATTC 2681
QY 4976 TTCTAATGATGATCGTACCGAGCTATCGTGGCGCTTGAATGCACTTTGATATC 5035
DB 2680 TTTGTTGCTGGATGACTGTTTCAAGCCTTCAGGCAATGATATCGTTACTTTAAATGA 2621
QY 5036 AAGCATGTTGCTTGGCGCTGCAAGTTCCTTGGGTGTTGTTCTATATGATGCTCCAT 5095
DB 2620 ACTGCGGCTTATTTGATGATGCTGTTGCGAGTATCTCTTATTTCAACTCAATAC 2561
QY 5096 ATGTCATGTTCTTGGTTGCGCGGATGTTACTTTCATTCATGATTCGCGCAGCAT 5155
DB 2560 ATGTCGCAATTTGACAGAACTATGCTATGTCATGTTGTTTCCATATCTC-----TTG 2507
QY 5156 GCTATGAGCTTTTACTGTTGTCGCGCAACGCGACATGATTCAGATGCAACGCTGCT 5215
DB 2506 ACTTCTCTTCCGAGAGGCTGCTCTTTTCAAAAACAGAGGCGATACCAACTTGCAG 2447
QY 5216 CCAAGTCTGACAGAACCAACCAAGCCGAGAACCGGAGAAATTTTCAAGAT 5275
DB 2446 GACAGATTTGTTGCTCAAGAAAGACAAATTTTGAAACATGAAACGTTGAATCT 2387
QY 5276 TCCACATATCCAGGCACTTTTGAACGCTGAAGCTATCGCATGAGCGCTCAGCAT 5335
DB 2386 TCGGTGAATTTATACGCCCACTAATGCGCAAGTGAAGATTTGATGATCAAGCAT 2327
QY 5336 GCCATGTTGCGACGGAAGCTTGGCTCAGGTGTCATGCTGCCATCTGTCAGCGTTTGA 5395
DB 2326 CTAATTTTGCATCAGGTGTCATGCGGCAAGGTCTAGTCATTTGAACCAAGCCAGGTGAG 2267
QY 5396 CTGTTTCAACAGTGAAGCGGAAGATTCGTGTCCTTCCCATCTGTCAGCGTTTGA 5455
DB 2266 TTGACCTTCAAGTTAATGAGACAGTACGCTTCTTCCCTTACAGACATGCAATGCGC 2207
QY 5456 GTCCGCACTAAGGCTGAGATGTTTCCATATGATATCTTGAATGACATGTTTTCAC 5515
DB 2206 ATTTGCTTGAAGAG-----GATTTGAATGCTCATTCACATCGATATGAT 2159
QY 5516 ACCGTAATCTCAACGCGACGCACTTTAACCCGCTGAAGAAAGCAGGCGATGAAGTCAA 5575
DB 2158 ACAGTATGCTTGAATGCGCAAGGTTTGAAGTCTTGAATGTCCAAGGAGTACGTTTCA 2099
QY 5576 GCAGGAGCTGCTGTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5635
DB 2098 GTTGTGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2039
QY 5636 ACCAGCGGATGTTGTTTGA 5658
DB 2038 GAACTCTGTTATCATCACCA 2016

```

RESULT 10
 US-09-583-110-463


```

; Sequence 463, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: P4H00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIORITY FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 463
; LENGTH: 1884
; TYPE: DNA
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-463

Query Match      2.8%; Score 167.8; DB 4; Length 1884;
Best Local Similarity 46.8%; Pred. No. 1,2e-41;
Matches 919; Conservative 0; Mismatches 962; Indels 84; Gaps 9;

QY 3779 ATGACCAATAGAGACTCGCGCAAGCATCTCGCGCACTTGGCGGCAACCAAGATGTG 3838
DB 1 ATGAACAAATCAGAAATTCGCAAAAAGATCATCGATCGCTTGGCGGCAAGTGAATGTC 60
QY 3839 GTGCGCGCGCACTGTGCAAGCGCTTTACGCGCTGCTCAAGACCAAGATGTG 3898
DB 61 AATAGGTGCGCACTGTGCAAGCGCTGTCTACGCTGTATGTCTCAAGATGAAGAAATC 120
QY 3899 GATCGCAAGTCTGTGATGATGATCCAGATCTGAAGGCACTTTGAACTGGCGCATG 3958
DB 121 AATTAAGAAGTATGAGAACTTGAAAGATTCAGAGTGTCTTTAACTCAGGGCAA 180
QY 3959 TTCCAGATCATGTGCGGCGCAGGCAATGTGATCATGTTTCAAGAACTTCATGACGA 4018
DB 181 TACCAAAATTAATCTTGTGATCAGGTACAGTAAACAAATGTACATGATGATTTGTACTT 240
QY 4019 ACCGCCAAGACATGCTGTGTCCACAGAGCACTCAAGATGTTGGCTTAACAAGGCC 4078
DB 241 GCGTTACCA-----CATCATTAAGATGACATTAAGACAGAACTTCTAAACAAGG 294
QY 4079 AACTGCTTACGCGTGTGTAAGGATTTGGCGGCAATTTGTCCTCCGTGATTCATTC 4138
DB 295 AACTGCTTACGCGTGTGTAAGGATTTGGCGGCAATTTGTCCTCCGTGATTCATTC 354
QY 4139 TTGCTGTGCGGCTGTCTCATGCTATCAACAATGTGTGGTGGCAGATCTGTTC 4198
DB 355 ATCGTAGCAGCAGGTCTCTTCAATGAGGTGTGCGGTCTTTCAACGCTTTGAATGCCA 414
QY 4199 GGTCCGCAATCATCTGTGTGAGATGTTCCTCATGATCAGAGGTGTGCTGAGATGATCAG 4258
DB 415 CTTTCCAG-----GTGACTTTGCAACTTACACACA 444
QY 4259 CTGATGCACTGTGCGCGTGTGCTTGTGCAAGTGTGTGTTGTTTCCAGCCAAACAAG 4318
DB 445 ATCTTGACAGATACAGCTTTCATCATCTTGCAGGTTTGTGTGTGTGTCACCTTCGCT 504
QY 4319 CGTTTCGTGTGCAATGAGTTCCTGTGCGCGCGGCAATGTGTGCGATGTGTTCACAC 4378
DB 505 GATATTGTGTGAAATCTCTGCGTGTGTATGTTCTTGTGTATGATGCTTCTCTGTGCTCA 564
QY 4379 CTGTTTAAAGGCTACAGACGTGCGCGCAATGACCGGCGGGAATGCAATGTGTGTC 4438
DB 565 CTTTCAAAAGCTTGGGCAAGTTC-----TCAAGGTGTGAACTTAACAGAGATGAAC 615
QY 4439 CTGTTTGTGTGATGTGTCTCAAGCTGTGTACAGAGGCAAGGCGTGTCTGTGTGTG 4498
DB 616 TCTTTGTGTTT---CATCCTGTGTGTGTGTGTGCAAGGTTCCGTTCTTCCAGCTTCAAC 672

```

```

QY 4499 GTCTCTTGAGATTCTGCAACGATTCAGAGATTCCTGCAACAAGCACTCATGCGCACTGCA 4558
DB 673 ATCGGGGTTGTGAGAGCTTAATTTGAAAAGCTGTCCGAAGGTGTTCAGATGTCAATT 732
QY 4559 GACTTCTGTATGACCCCAAGTGTGTGACTGTGTCTACACCGGCTTCTTACGTTCAATGCT 4618
DB 733 GACCTTGTGTAAACACANTCGTGAACCTTTTGTGTATGTCTATCTTGTGACTCTTGTTC 792
QY 4619 ATTTGTTCAGCAATGCGCTGGGTGTGACTGTGTGCAACAGGTGTGCAAGGACTTAT 4678
DB 793 ATTTGACCGAGTTTCCACGCTGTGTGAAAATCACTCATTTATGTCTAACAAGCAATCTT 852
QY 4679 GATTTGCGGTGTGCAAGTGTGCGGCTGTGCTTTTGTGTGTGTGTGTGTGTGTGTGTGT 4738
DB 853 AGCATTTCAATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 912
QY 4739 ACTGCTGTGCAACAGTCTTCCGCAATTTGAGTGTGAGCTGTGCAACAGG---GTGGA 4795
DB 913 TCAAGTGTGCAACCATCTTCAACTTGTGTGAGTGTGCAATTTACTGTGTGTGATCATGCT 972
QY 4796 TCTTTCATCTTGTGCAACCGCATTCATGCGCAATTTGCGAGGTTGCAAGATGTTTGGCA 4855
DB 973 AACCATTCACAGCTATCATCACAGCTGTGTGACAGCTCAAGGTGTGTGTGTGTGTGTG 1032
QY 4856 GTGTCTTCTGTGCGCAAGTGAAGTGAAGTCAAGGCGCTTGTGAGTGTGTGTGTGTGTGT 4915
DB 1033 GTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1092
QY 4916 GCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4975
DB 1093 GCTTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1152
QY 4976 TTCTACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5035
DB 1153 TCTTCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1212
QY 5036 AAGGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5095
DB 1213 GCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1272
QY 5096 ATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5155
DB 1273 CAATTTCAACAATACCTTCTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1323
QY 5156 GCTTATGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5215
DB 1324 GCTTATCTTACATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1374
QY 5216 CGATGCTGTGCAAGACCAAGCCAGACAGACAGACAGACAGACAGACAGACAGACAGAC 5275
DB 1375 GCAAAACAAAGCTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1434
QY 5276 TCCACATCATTCAGAGCACTTGTGACCGGTGAAGTGTGTGTGTGTGTGTGTGTGTGT 5335
DB 1435 GAAACACTTGT---ACTTCTATGCTCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1491
QY 5336 GCGATGTTTGTGCAAGGAAAGCTTGTGCTCAAGTGTGTGTGTGTGTGTGTGTGTGTGT 5395
DB 1492 CCAAGCTTCTGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1551
QY 5396 CTGCTTTCACAGTGAAGGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5455
DB 1552 GTTATGTGACCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1611
QY 5456 GTCCGACATAAGGCTGAAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5515
DB 1612 TTGAAAACAAAGAAATG-----GTGCTGAAGTTTGTATTTCAAGCTTGTGTGT 1659
QY 5516 ACCGTAAACCTCAACGCAAGCACTTAAACCGCTGAAGAAAGCAAGGCGGCAATGAATCAA 5575
DB 1660 ACTGTATCTATGAAACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1719

```

QY 5576 GCAGGGAGCTGCTGTGTAATTCATATGATGCAATTAAGCTGCAAGTTATGAGTA 5635
DB 1720 GCTGCGCATGTTCTTGGAACATTTGACCAAAATCGCTGAGCTGAGCTTATGAT 1779
QY 5636 ACCAGCGCATGTTGTTGTAATTAACAAGAAACCGAGCTGTAAACCTTACGGTTTG 5695
DB 1780 ACAACATGTTATGTTATGTTAATAATACAGTGACTACGCTTCACTGCTCCAGTCCACA 1839
QY 5696 GCGGAATTGAAGCGGAGCCAACTGCTCAACGTCGCAAGAA 5740
DB 1840 GGTTCAGTTGCGAAGGGGATGCTGTGATCGAAGTGAATCTAA 1884
RESULT 11
US-09-107-433-647
Sequence 647, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNOSTIC
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 647:
SEQUENCE CHARACTERISTICS:
LENGTH: 1932 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
MOLECULE TYPE: DNA (genomic)
TOPOLOGY: circular
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...1932
SEQUENCE DESCRIPTION: SEQ ID NO: 647:
US-09-107-433-647
Query Match
Best Local Similarity 2.8%; Score 167.8; DB 4; Length 1932;
Matches 919; Conservative 0; Mismatches 962; Indels 84; Gaps 9;
QY 3779 ATGACCATTAAGACCTCGCGCAACGCATCTGCGGACATTGGCGGCAAGCAACATT 3838
|||||

DB 49 ATGAACATCAGAAATTGCAAAAAAGTCATGATGCTTGGCGGACGTGAAATGTC 108
QY 3839 GTCCGCCCGCACTGTGCAACGCCGTTTACGCTCGTGTCTCAAGACCAAGATGTG 3898
DB 109 AATAGTGTGGCCACTGTGCACTGCTTACGTGTCAATGTCGCAAGATGAAGAAAAATC 168
QY 3899 GATGCCCAAGTCTGAGATGATGATCCAGATCTGAAGGCACTTTGAAATCGCGGCAATG 3958
DB 169 AATTAAGATGATTTGAGAACTTGAAAAAGTTCAAGGTCTTTCTTTACTCAGGGCA 228
QY 3959 TTCCAGATCATGTCGGCCACAGCGAGTGTGATCATGTTTCAAGAACTCGATGACGA 4018
DB 229 TACCAATATCTTGTGTAAGGTACAGTTAACAAAAATGATGATGAAGTGTGTACTT 288
QY 4019 ACCTCCAAAGATGCTGTGTCCACAGAGAGCTCAAAAGATGTTGTGGCTTAACAGCC 4078
DB 289 GAGTTACCA-----CATCATCTAAGATGATGATGAAGAGAAAGTTGTAACAAAGG 342
QY 4079 AACTGTTCAAGCCGCTGTGAGATATTTGCGGACATTTTCGTCGCCGTGATTCATC 4138
DB 343 AACTGTTCAAGCTGTGATCCGTACTTGGTGTATGTTTGTTCATCATCCAGTT 402
QY 4139 TTGTTGTGCGGCTGTGCTCATGCTATCAACAATGTTGTGCGCAGATCTGTTTC 4198
DB 403 ATCGTAGCAGAGGCTCTTCATGGGTGTGCGGTCTTTCAACGCTCTTGAATGCA 462
QY 4199 GCTCCGCAATCATGCTGTGAATGATGTCCTCAGATCAGCGGTGTCGATGATATCAC 4258
DB 463 CTTCAG-----GTGACTTGCACATTTACACAA 492
QY 4259 CTGATGCAATCGCGCGCTTGCGGCTTTCGCAAGTGTGTTGTTTACCGCAACCAAG 4318
DB 493 ATCTTGAAAGATCAGCTTTCATCATCTTGCAAGTTTGTGTGTGCTCAACCTCGT 552
QY 4319 CGTTTCGATGCAATGATCTCTGCGCGCGCATGTGTATGCGATGCTGTCCAAAC 4378
DB 553 GTATTGTGTGAAATCTGCGGTGTGTATCTTCTGTGTATGATGCTGTCTGCTCA 612
QY 4379 CTGTTAAGCGCTACGATGCTGCGCGCAACATGACCGGCGGCAATGCAATGCTGC 4438
DB 613 CTTCAAACGCTTGGGCAAGTGC-----TCAAGTGTGAAATGAAACGATGAC 663
QY 4439 CTGTTGTGTTGATGTTGCTCAAGCTGTACAGGCAACCGCTTCTGTGCTGTG 4498
DB 664 TTCTTTGTTT---CATCCGTGTTGTGTTTGAAGTTTCGTTCTTCAGCCTTATC 720
QY 4499 GTCTCTGATTTCTGCAACGATGCAAGATTTCTGCAAGCAACGACTATGCGCACTGCA 4558
DB 721 ATCGGGTGTGTGAGCTAAATTTGAAAAAGCTGTCCGCAAGTGTGTCCAGATGCTATT 780
QY 4559 GACTTCGTGATCACCCAGTGTGATCTGCTGCTCACCGGCTTCTTACGTTCAATGCT 4618
DB 781 GACTCTTGTGTAACCAATGTGTGACACTTTTGTGTCAATGCTTATCTTGTGACTTTTGTG 840
QY 4619 ATTGTTCAGCAATGCGTGTGAGTGTGACTTGTGCAACAGCTCTGCAAGGACTAT 4678
DB 841 ATTGAACAGTTTTCAGTTGTTGAAAAACTAATCCTTATGTGTAAGAAAGCAATCTT 900
QY 4679 GATTTGCGTGTGCAAGTGTGCGGTGCTTTCGTTGTGCTGTATCTCAACATGCTATT 4738
DB 901 AGCATTTCAATTTGTGTTGTGTGTTTCTTGAATGTGTGGGTTCACAAATGATGCTGTG 960
QY 4739 ACTGTCGACCAATCTCTCCCGCAATGAGCTGAGAGCTGTTCAACAGG---GTGGA 4795
DB 961 TAGGTGTGACCAATCTTCACTTGCTTGAAGTCAATTAAGTCTGCTGACATGCT 1020
QY 4796 TCCTTCATCTTGCACACGATTCATGCGCAATATGCGCAGGCTGAGCAATGTTTGCA 4855
DB 1021 AACCATTAACGATATATACACAGCTGTATGACAGCTCAAGGCTGTACTGTTCG 1080
QY 4856 GTTCTTCTCTAGCAAGATGAAGAAAGCTCAAGGCTCTTGAAGTGTCTTCAAGTGTCTCC 4915
DB 1081 GTTGTGTTAACAAAAATCCAAAGCTGAACACTTGTCTTTCGCGGCTGTCTTCT 1140
|||||

```

QY 4916 GCGTTCCTGGTATTACAGAGCTGGATCTTGGGTGAACCTTGCCGCTGGCGG 4975
Db 1141 GCGTTCCTGGTATTACAGAGCTGGATCTTGGGTGAACCTTGCCGCTGGCGG 1200
QY 4976 TTCTACATGTGATCCGATCCGAGCTATCGGTGGCGCTTGTGATTCACCTTGTGATTC 5035
Db 1201 TTCTTCCTTCAATGATGTGGTGAATCGGTGGGTGATGTGGCTTCTATCTTGGACCT 1260
QY 5036 AAGCAGTTGCGTGGCGCTGAGGTTCTTGGGTGTTCTATGATGCTCCAGAT 5095
Db 1261 GCGTGTACTGTGATGATGATCAACATCCCTGTGATCAATGCTTATGTGGTACGGA 1320
QY 5096 ATGGCATGTCTGTGGTTCGCGGTGATTAACCTTGTGATCCATTCGCGCGAGGAT 5155
Db 1321 CAATTCACAAATACCTTCTTATGTGATGCTGAT-----CATTTGCCCTGGTGT 1371
QY 5156 GCTTATGCGCTTACTTGTGTGCGCGCAACGCGACATTCATCCAGATCCACCGCTGCT 5215
Db 1372 GCTTACTTACATGTTGTGTTACGAAG-----TGAAGTACCGCACTGCAAGCT 1422
QY 5216 CCAATGCTGCAAGAACGACCAAGCCGACGAGAACACCGCGAATTTTCAACGAT 5275
Db 1423 GCAAAACAAAGCTGAAGTGGCTGAAGAAAGAAAGATTGCGCGACAGCTCTTCAAAAT 1482
QY 5276 TCCACATCATCCAGGACCTTGTGACCGGTGAAGCTATGCACTGACAGCGTCAAGCAT 5335
Db 1483 GAAACCTGTA--ACTCTTATCGTGGGATGTGTGCTTGTGATGTCATGATGAC 1539
QY 5336 GCCATGTTTCCAGCGGAAAGCTTGGCTGAGTGTGCGATCGTCCCAACGAAGGGCGAG 5395
Db 1540 CCAATGTTTCAAGTGAAGCTATGGAACAAAGATGCTGTGAACCAAGCGCAAGCGCG 1599
QY 5396 CTGGTTTCAACAGTGAAGCGAAAGATCGTGTGGCGCTTCCATCTGTCAAGCTTTGCA 5455
Db 1600 GTCTATGACACAGCTGATCTGAATTCATTAATGCTTCCAAACAGGCGACGCTTTGGT 1659
QY 5456 GTCCGACCTAAGGCTGAGATGTTCCATGTCGATGTCGATGTCGATGTCGATGTCGAT 5515
Db 1660 TTGAACCAAGAAATG-----GTGCTGAAGTTTGAATGATGCTGTGATGTCGAT 1707
QY 5516 ACCGTAACCTCAACGCGACGACCTTAAACCGCTGAAGAGAGGCGCATGATGATCAAA 5575
Db 1708 ACTGATATCATGACGCTGACGCTTTTGAAGCAAAAGTGTCTCAAGTAAATAGTGAAA 1767
QY 5576 GCAAGGAGCTGCTGTGATTCGATGATGATGATGATGATGATGATGATGATGATGAT 5635
Db 1768 GCTGGGATGTTCTTGAACATTTGACTCAACAAATGCTGACGCTGACCTGATGATGAT 1827
QY 5636 ACCAGCGCATTTGTTTTCGATTCAGAAACCCGACCTGTAAACCTTAACGCTTTG 5695
Db 1828 ACAACCAATGTTATGCTTCAATATACAGCTGATGATGATGATGATGATGATGATGAT 1887
QY 5696 GCGCAATTTGAAGCGGAGCAACCTGCTCAACGTCGCAAAAGAA 5740
Db 1888 GGTTCAGTTGCGAAGGCGGATGCTGTGATGATGATGATGATGATGATGATGATGAT 1932

```

```

RESULT 12
US-09-489-039A-1806
; Sequence 1806, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 1806

```

```

; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-1806

Query Match
Best Local Similarity 47.8%; Pred. No. 5.8e-40;
Matches 612; Conservative 0; Mismatches 643; Indels 24; Gaps 4;

QY 3819 TTGGCGGCGAGAACCAATTTGCGCGCGCACACTGTGCAACGCTTTACCGCTGTCG 3878
Db 56 TGGGCGGCGCGAGTAACTGCGACCGTCAAGCTGATTAACCGCTTGTGTCG 115
QY 3879 TGAAGACACCAAGATGATTCGCAAAAGTGTGATGATGATGATGATGATGATGATGATGAT 3938
Db 116 TGAAGACCGCGCATCGCGACAGCAAAAGATGATGATGATGATGATGATGATGATGATGATGAT 175
QY 3939 CTTTGAACCTGCGCGCATGTTTCAGATCATCGTGGCGCAGCGCATGATGATGATGATGAT 3998
Db 176 GCTTCACTAAGCGCGCGCATGTTTCAGATCATCGTGGCGCAGCGCATGATGATGATGATGATGAT 230
QY 3999 TCAAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4058
Db 231 -CTAAGAGCCCTGCTGGCGACACCGCGCAGACCTCCGCGCAAGAGAGAGTCAACG 289
QY 4059 ATGTTGTGCTTAAACCGCACTGTTTCAAGCGGTGCTGATGATGATGATGATGATGATGATGAT 4118
Db 290 AAGCGCGCGCGCAAAATATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 349
QY 4119 TGTCCCGCTGATTCATCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 4178
Db 350 TCTTTCCTGCTGCTGCGCGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 409
QY 4179 TGTGTCGAGATCTGTTGCTGCGCAATCACTGTTGATGATGATGATGATGATGATGATGATGAT 4238
Db 410 TGGGAGACCTGCGCATGAGCAATGAGACGCTGCGCAAGATGATGATGATGATGATGATGATGATGAT 469
QY 4239 GTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4298
Db 470 CCATCTATGATCTTCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 529
QY 4299 TTGTTTTCACCGCAACCAAGCGTTTTCGTTGATGATGATGATGATGATGATGATGATGATGAT 4358
Db 530 TCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 589
QY 4359 TGGGATGTTTCCCAACCTGTTTAAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 4418
Db 590 TCAACCTGCTTCCCGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 649
QY 4419 GCGAAATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4478
Db 650 -----TGTGAACTTCCGCTGTTTAACTGATGATGATGATGATGATGATGATGATGATGATGAT 697
QY 4479 CCGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4538
Db 698 AGTATATCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 757
QY 4539 AGGACTATGCGGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4598
Db 818 TGTTCCTGCGAGCGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 877
QY 4599 GCTTCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCT 4658
Db 878 TGTTCCTGCGAGCGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 937
QY 4716 TGTCTACTACCAATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 4775
Db 938 TCTCTACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997

```



```

QY 4916 GCTGTTCTGTATTTACAGAGCTGCGATCTTCGATGTAACCTTGCCCTGCGCGCCG 4975
DB 8619 GCTTCTTGTATTTACAGAGCTGCTATCTTCGCGGTAACTTGCGCTTCCGTAAACCA 8560
QY 4976 TTCTACATTTGATTCGTAACCGAGCTATCGGTGGCTTTGATTCACCTTTTGTATATC 5035
DB 8559 TTCTTCTTCAATTTGATTCGTAACCGAGCTATCGGTGGCTTTGATTCACCTTTTGTATATC 8500
QY 5036 AAGCAGTTGCGTGGCGCTGAGCTTCTTGCGGTGTTTCTATTTATGCTCCAGAT 5095
DB 8499 GCTGTACTGTAATGTAATCAACATCATCCCTGTACAAATCTTATTTGTATACGA 8440
QY 5096 ATGTATGTTCTTGTGTTGCGGATTTACCTTTGATGATGATGCGGCGAGGAT 5155
DB 8439 CAATTTCCACATACCTTTCTTATGTAAGCTAT-----CATTTCCCTTGTTTT 8389
QY 5156 GCTTATGCGCTTATCTTGTTGCGCGCAACGCGAGCATTTATCAGATCAACCGCTGCT 5215
DB 8388 GCTTCTTACTTACATGTTTGTACGAAGA-----TGAGTACGCAACTGAGCT 8338
QY 5216 CAGAGCTGCGAGAACGACCAAGCCGAGCAAGAACACCCGAGATTTTCAAAAGAT 5275
DB 8337 GCAGAACGAGTGAAGTGTGTAAGAAAGAAAGAGTGGCCGACGACCTCTTCAAAAT 8278
QY 5276 TCCACCATGATCCAGGACCTTTGACCGGTGAAGCTATGCACTGAGAGCGTCAAGCAT 5335
DB 8277 GAAACCTGTGA---ACTCTATCTGTGGTGAATGTTGCTCTTGTGATGATGATGAC 8221
QY 5336 GCCATGTTTCCAGCGGAAAGCTGAGGTGTTGCGATGATGCTCCGACCAAGGCGCAG 5395
DB 8220 CCAAGCTTCTCAAGTGAAGCTATGGAACAAGTATCTGTTGAAACCAAGCCAGGCGTG 8161
QY 5396 CTGTTTCAACAGTGAAGGAAAGATGCTGCGCTTCCATCTGTCAGCCTTTGCGA 5455
DB 8160 GTCTATGACGACGAGCTGATCTGAAGTTCAATTTGCTTCAACAGGCGACCTTTGGT 8101
QY 5456 GTCCGACATGAAGGCTGAAGATGTTCCATATGTAATCTTATGATGACATTTGTTGAC 5515
DB 8100 TTGAAGCAAGAAAGT-----GTGCTGAAGTTTATGATGATGATGAT 8053
QY 5516 ACCGTAACCTCAACGCGACGATTTAACCCGCTGAGAACAGAGGCGATGAAGTCAA 5575
DB 8052 ACTGATATCATGAACGCTGATGCTTTTGAAGAAAGTGTCTCAAGTAAATGAGTGAAA 7993
QY 5576 GCGAGGAGAGCTGCTGTGTAATTCATATTTATGATGATGATGATGATGATGATGAT 5635
DB 7992 GCTGCGATGTTCTTGAACATTTGATCTCAACAAATGCTGCGAGCTGATGATGAT 7933
QY 5636 ACCAGCGGATGTTGTTTGAATTAACAAGAAACCGACCTGTAACCTTAAGCTTTG 5695
DB 7932 ACAACAAATGTTATGTTAACAATACAGGTGATGATGATGATGATGATGATGATGAT 7873
QY 5696 GCGCAAAATGAAGCGGAGCAACCTGCTCAAGCTGCGCAAGAAA 5740
DB 7872 GGTTCAGTTCTAAGGCGGATGCTGTGATGATGATGATGATGATGATGATGATGAT 7828

```

```

RESULT 14
US-09-134-001C-1277
; Sequence 1277, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14

```

```

; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1277
; LENGTH: 732
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (20) (22) (32)
; OTHER INFORMATION: Identity of nucleotide at the above locations are unknown.
US-09-134-001C-1277

```

```

Query Match 2.3%; Score 139; DB 3; Length 732;
Best Local Similarity 56.6%; Pred. No. 7.9e-33;
Matches 306; Conservative 0; Mismatches 220; Indels 15; Gaps 2;

```

```

QY 4569 TCACCCAGTGTGACTGCTGCTACCGGCTTCTTACGTTCAATGCTATGCTCAG 4628
DB 80 TGACCCCAATTTATCAATTTTATTACAGCATTTATTAACATTTTATTGAGGCGCTG 139
QY 4629 CAATGCGTGGGTGAGTCTGCTGACACGCTGCGAGGACTGATGATTTGCGTG 4688
DB 140 TCACCTGCAATTTAGTATTTGTTATCTGATGATGATGATGATGATGATGATGATGAT 199
QY 4689 GTCCAGTGGCGGTGCTGCTTTCGCTGCTACTCAACATGCTTATCACTGCTGTC 4748
DB 200 GAGCTATGTTGAGCTTATTTGTTTATTTATATGCGCCATGCTCATACAGAAATGC 259
QY 4749 ACCAGCTCTCCCGCAATGAGCTG-----GAGCTGTTCAACGAGGTGAT 4796
DB 260 ACCATATCTTTATTTAGATTTGAAACATTAATTTGATGATGATGATGATGATGATGATGAT 319
QY 4797 CCTTATCTTGCAACCGCATTCATGCGCATATGCGGAGGTGACAGATTTGGCAG 4856
DB 320 CATTTATCTTCCCATCCCAACAGATGCAAAATTTGCAACAGGTGTCAGCTTTAGCTG 379
QY 4857 TGTCT---TCTACGAAAGTGAAGATCTCAAGGCGCTTGCAGTGTCT 4913
DB 380 CATTTTATATCTTATGAGCAAAATTAATAATTAAGTGTCTTCCGCGGCGGTATTT 439
QY 4914 CCGCTGTTCTGTAATTAACAGAGCTGCGATCTTGTGTAACCTTGCGCTGCGCTGCG 4973
DB 440 CAGCTTACTAGATTAACAGAACGAAATGTTTGTGATCACTTTAATGATGATATC 499
QY 4974 GCTTATCAATTTGATGCTGATCCGAGCTATGCGGTGCTTTGATTTGATGATGATGATGAT 5033
DB 500 CATTTATGAGTGTGCTTCAAGATCAAGTATAGGTGCGCTTATTTTCAATTTCTTCAAG 559
QY 5034 TCAAGGAGTGGCGTGGGCGCTGCGAGTTCTTGCGGTGTTGTTCTTATGATGATGATGAT 5093
DB 560 TAAAGCGATAGCGCTTGTACAGTGAATTAAGTATTTATTTATTTATTTATTTATTTAT 619
QY 5094 A 5094
DB 620 A 620

```

```

RESULT 15
US-09-107-532A-378
; Sequence 378, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSER: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: PC
 OPERATING SYSTEM: <Unknown>
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/107,532A
 FILING DATE: 30-Jun-1998
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/085,598
 FILING DATE: 14 May 1998
 APPLICATION NUMBER: 60/051571
 FILING DATE: July 2, 1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Arinfield, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781)893-5007
 TELEFAX: (781)893-8277
 INFORMATION FOR SEQ ID NO: 378:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 714 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: circular
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (8) LOCATION 1..714
 SEQUENCE DESCRIPTION: SEQ ID NO: 378:
 US-09-107-532A-378

Query Match 2.1%; Score 124.8; DB 4; Length 714;
 Best Local Similarity 52.5%; Pred. No. 2.5e-28;
 Matches 380; Conservative 0; Mismatches 322; Indels 22; Gaps 4;

QY 1544 CACCATGGAATCTCTATCTGCAAGAGCAAGAGTGG--CAAGCAGTTGCACT 1600
 DB 9 CACTATGGAATCATCTCGTAAAAACCTGAGAGCGGAAAAAGCATTTGAAAT 68
 QY 1601 CCTATCGACCCCTCGCCAAAGGCTGAACTTGGGGCTTGCACAGATCCTCAAC 1660
 DB 69 CATCAAGAGGTATGAAAGAGTGCAGAAAGTATTAAGATTGGCAACAGAGCACACC 128
 QY 1661 ACTGAGTACTACCAAGAGCTCATTCGATGTATGAGCTGGGAGAGTGCATTCAGAA 1720
 DB 129 AGTTACTTATATAAGAAATGACTTCAA-----CGATTTGATTTCACTGA 176
 QY 1721 CTGCAAGGCAATCTGTGTGATGAATACGTGGACTAACCCGTGACATGAAGAGCTA 1780
 DB 177 AATGACTTCTGAACCTAGATGAGTACCTGACTAGCGGCTTCAGATGAACAAAGCTA 236
 QY 1781 CTTTAAACCATTCGAAAGAGTCACTGACACATCGACATCGATTGATGAAGAGTCTA 1840
 DB 237 CGGTCACTTTTCAATGATGTCCTCAATATTTGATTAATAAACCATTTCA-----AAGAAACGTT 290
 QY 1841 CAGCCAGATGGTGCAAAACCTGATCCATACAGAGAGCTGCAGAGTATAGAGCAAGAT 1900
 DB 291 TGTACCAAGATGAAAAAGCAAGAGATCTGAGACAGATCAAAAAGTATGATCAATCAT 350
 QY 1901 CGCTGCAGAATCCGTTGAAGTTCAAAATCTTGGCATCGCGGAAACGG--CACATGGCTTT 1959
 DB 351 CGATGCACATCCAAATCGATATCCAAATCTTAGAAATCGGGCAAAATGACATATCGGTTT 410
 QY 1960 CATTAACCATCATCTCTCTGTCAGAGTGAACAAAGGTCAAGGCGCTGCACCTATAAAC 2019
 DB 411 CATTAACCTGGTGCTTCTTTGAGCGAAGCACTTCACTGATGATTGAATTAACGAATCAAC 470
 QY 2020 TGTGAGGACAAAGCTGATCTTCAACACATCGAAGAGTCCACCCACGCGTCAC 2079

DB 471 AATCAATCAAAACAAACGATTTCTGATTAAGCAGAGATGTACCAACAAAGCTGTTTC 530
 QY 2080 CCAGGTTTGGGCACTTTGTCCCGCGCAAAACATGCTGTTGGTGCACACTGGTGAAG 2139
 DB 531 AATGGAAATCGGTTCAATTATGAAGAAATGCTTGTGATGGCATTTGGCGAAAG 590
 QY 2140 AAAGCGAGCGCATTCGCGGGAAGTGTGAAGAGGCGGAGTACTGCTTCTGCCAGGTTTC 2199
 DB 591 CAAGCTGATCAATCAAGGAGATGATTGAAGACCTGTAAAGCATCTTTGCCAGCAAG 650
 QY 2200 CATCTGTAGTGCACACATATGCCACATCATCTGTGATGAAGAGAGCATATCCAAAGCT 2259
 DB 651 TGTATGCAAAATCACGACATATGTTGTGTGATCATGATGAAGAGCAGCAAGTAAGTT 710
 QY 2260 GGAA 2263
 DB 711 GTAA 714

Search completed: March 9, 2005, 13:51:46
 Job time : 960 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 9, 2005, 12:11:23 ; Search time 2980 Seconds

(without alignments)
11885.615 Million cell updates/sec

Title: US-10-019-284B-1

Perfect score: 5969
Sequence: 1 agtcgcgtcagccaccatc.....tgatatttcgcgcctgaa 5969

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5401638 seqs, 2966923429 residues

Total number of hits satisfying chosen parameters: 10803276

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09C_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10F_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	5656.2	94.8	3309400	9	US-09-738-626-1
2	1874.2	31.4	1963	9	US-09-738-626-2904
3	1412.6	23.7	1527	17	US-10-450-055-9
4	1321.2	22.1	1342	18	US-10-781-014-347
5	1267.8	21.2	1299	9	US-09-738-626-2905
6	1237.4	20.7	1287	18	US-10-781-014-413
7	1237.4	20.7	1287	18	US-10-781-014-415
8	1110.4	18.6	1152	9	US-09-738-626-2907
9	762	11.3	759	9	US-10-781-014-423
10	671.8	11.3	759	9	US-09-738-626-2906
11	294	4.9	1543	19	US-10-494-541-9

ALIGNMENTS

RESULT 1

US-09-738-626-1/c

Sequence 1, Application US/09738626

Publication No. US20020197605A1

GENERAL INFORMATION:

APPLICANT: NAKAGAWA, SATOSHI

APPLICANT: MIZOGUCHI, HIROSHI

APPLICANT: ANDO, SEIKO

APPLICANT: HAYASHI, MIKIRO

APPLICANT: OCHIAI, KEIKO

APPLICANT: YOKOI, HARUHIKO

APPLICANT: TATEISHI, NAKO

APPLICANT: SENOH, AKIHIRO

APPLICANT: IKEDA, MASATO

APPLICANT: OZAKI, AKIO

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-125

CURRENT APPLICATION NUMBER: US/09/738, 626

PRIOR FILING DATE: 2000-12-18

PRIOR APPLICATION NUMBER: JP 99/377484

PRIOR FILING DATE: 1999-12-16

PRIOR APPLICATION NUMBER: JP 00/159162

PRIOR FILING DATE: 2000-04-07

PRIOR APPLICATION NUMBER: JP 00/280988

PRIOR FILING DATE: 2000-08-03

NUMBER OF SEQ ID NOS: 7059

SOFTWARE: Patent ver. 3.0

SEQ ID NO 1

LENGTH: 3309400

TYPE: DNA

ORGANISM: Corynebacterium glutamicum

US-09-738-626-1

Query Match 94.8%; Score 5656.2; DB 9; Length 3309400;

C	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
	215.2	208	192.6	192	179.2	177.6	177.6	174.2	173.8	173.8	172.2	172.2	172.2	163.6	161	154.2	154.2	153	139.2	134	133.6	133.6	122.2	122.2	121.2	121.2	118.4	117.4	117.4	111.4	111.4	109.8	107.2	
	3.6	3.5	3.2	3.2	3.0	3.0	3.0	2.9	2.9	2.9	2.9	2.9	2.9	2.7	2.6	2.6	2.6	2.6	2.3	2.2	2.2	2.2	2.0	2.0	2.0	2.0	2.0	2.0	1.9	1.9	1.8	1.8	1.8	
	1059	777	1212	975	3885	30246	30246	1380	2127	2162598	1965	3895	3895	1884	9769	9769	9769	5840	3037	1137	684707	3011208	567	1143	9025608	699	936	783	1185	1098	116020	3011208	1014	798
	US-09-746-660A-11	US-10-282-122A-17961	US-10-282-122A-17959	US-09-974-300-829	US-10-857-625-2	US-08-781-986A-56	US-10-329-624-56	US-09-974-300-685	US-10-474-776-150	US-10-472-928-4979	US-10-472-928-3911	US-08-961-527-201	US-10-158-844-201	US-09-815-242-9369	US-09-738-626-2903	US-08-961-527-30	US-10-158-844-30	US-09-070-927A-31	US-10-398-221-3599	US-10-282-122A-34025	US-10-398-221-2058	US-10-398-221-9	US-09-070-927A-664	US-10-156-761-3925	US-10-156-761-1	US-10-282-122A-21630	US-09-738-626-2908	US-10-156-761-3005	US-10-474-300-734	US-09-974-300-734	US-10-398-221-10	US-10-398-221-2058	US-09-974-300-822	US-10-282-122A-19435
	Sequence 31, Appl	Sequence 17961, A	Sequence 17959, A	Sequence 829, App	Sequence 2, Appl	Sequence 56, Appl	Sequence 56, Appl	Sequence 685, App	Sequence 160, App	Sequence 4979, App	Sequence 3911, App	Sequence 201, App	Sequence 9369, App	Sequence 2903, App	Sequence 30, Appl	Sequence 30, Appl	Sequence 31, Appl	Sequence 3599, App	Sequence 34025, A	Sequence 2058, App	Sequence 664, App	Sequence 3925, App	Sequence 1, Appl	Sequence 21630, A	Sequence 2908, App	Sequence 3005, App	Sequence 152, App	Sequence 734, App	Sequence 10, Appl	Sequence 2058, App	Sequence 822, App	Sequence 19435, A		

Best Local Similarity 97.1%; Pred. No. 0;
Matches 5793; Conservative 0; Mismatches 173; Indels 3; Gaps 3;

```
QY 1 AGTCGTCGACGCGCACATTGATGTGTGTCACCGAGCTTGGCGAGGCTTTCTACATCT 60
Db 2815585 AGTCGTCGACGCGCACATTGATGTGTGTCACCGAGCTTGGCGAGGCTTTCTACATCC 2815526
QY 61 AGGCTCCGCGCGCGGTCGATGAGGTCATTACCGGTGGGATGACCGCGGTGAAGTTGG 120
Db 2815525 ACACCTCCGTCGCGCGGTCGATGAGGTCATTACCGGTGGGATGACCGCGGTGAAGTTGG 2815466
QY 121 GAACCATGTGTCTCTGTGGGTTGAGGAAACGAGTGGCGGGTGAAGATTTTCAAGT 180
Db 2815465 GAACCATGTGTCTCTGTGGGTTGAGGAAACGAGTGGCGGGTGAAGATTTTCAAGT 2815406
QY 181 TCTGCAGTTTTTAAGTTATGATCATACGCTTGAAGGCTGAGTTAATTCAGTAGACTG 240
Db 2815405 TCTGCAGTTTTTAAGTTATGATCATACGCTTGAAGGCTGAGTTAATTCAGTAGACTG 2815346
QY 241 CAACACGAGCCCTCAAGTCGGAAGATTAATTAACCTGATCCGTTAGCATTAAGACATCA 300
Db 2815345 CAACACGAGCCCTCAAGTCGGAAGATTAATTAACCTGATCCGTTAGCATTAAGACATCA 2815286
QY 301 CGTCCTATGCTTGTGGAAGAAACCAATTAACCTCAGAAAGATGCGAAGTGTGCTATT 360
Db 2815285 CGTCCTATGCTTGTGGAAGAAACCAATTAACCTCAGAAAGATGCGAAGTGTGCTATT 2815226
QY 361 ATCAGAAAAATGCAAGGTCAAGCATTTAAAAAATTGAGGAAAGATTTTCCCCCTCG 420
Db 2815225 ATCAGAAAAATGCAAGGTCAAGCATTTAAAAAATTGAGGAAAGATTTTCCCCCTCG 2815166
QY 421 GGGTATGATGGCTTCTCCAACTCGAAAAAGGCACTCATCGGAACCTCTCGAGAAC 480
Db 2815165 GGGTATGATGGCTTCTCTCCAACTCGAAAAAGGCACTCATCGGAACCTCTCGAGAAC 2815106
QY 481 CAGCAGCTTAAAAACGAGGATTCACCCCGAACTCCCAAGATTGTTCCCGGTTTATATG 540
Db 2815105 CAGCAGCTTAAAAACGAGGATTCACCCCGAACTCCCAAGATTGTTCCCGGTTTATATG 2815046
QY 541 ATCTTATATATCAGCGTGGAAACGGTGGCGCTTTCTACGGGAACGCGACGACGCGA 600
Db 2815045 ATCTTATATATCAGCGTGGAAACGGTGGCGCTTTCTACGGGAACGCGACGACGCGA 2814986
QY 601 GGAACACGCGGAGTATCACCGGAAACATGCGACGACGCGTGAATGTTGCAACATGTT 660
Db 2814985 GGAACACGCGGAGTATCACCGGAAACATGCGACGACGCGTGAATGTTGCAACATGTT 2814926
QY 661 CGGCGCGGCTGACGCACTGGGACGCGAGGTGAAAACTTATTCCTTGTGTGAAGAG 720
Db 2814925 CGGCGCGGCTGACGCACTGGGACGCGAGGTGAAAACTTATTCCTTGTGTGAAGAG 2814866
QY 721 TCCTGTGTGCGGCACTTCACTCGAGGGCCCTTTATCAACGACATGCCCTTGTGTGCT 780
Db 2814865 GCCTGTGTGCGGCACTTCACTCGAGGGCTTTTATCAACGACATGCCCTTGTGTGCT 2814806
QY 781 AAAACCCGGAATTTATTTTCCCGGCAACCAACAGATCTTCCCGGGTATCCATGCGG 840
Db 2814805 AAAACCCGGAATTTATTTTCCCGGCAACCAACAGATCTTCCCGGGTATCCATGCGG 2814746
QY 841 GAAAAAGTTGATCAATTCGATCAAGTAGCGCCGGAACCTGACATCTTTCTAGCTTC 900
Db 2814745 GAAAAAGTTGATCAATTCGATCAAGTAGCGCCGGAACCTGACATCTTTCTAGCTTC 2814686
QY 901 TCGATCTCTGCGGACGCGACATCATTTGCTTCTTGGGCACTGTATGAGATTTTG 960
Db 2814685 TCGATCTCTGCGGACGCGACATCATTTGCTTCTTGGGCACTGTATGAGATTTTG 2814626
QY 961 ATACCACTACAGCGCAATTTGCTTGGCTTAAGAAAAATGTAACGCGTACAGGCTAGCC 1020
Db 2814625 ATACCACTACAGCGCAATTTGCTTGGCTTAAGAAAAATGTAACGCGTACAGGCTAGCC 2814566
QY 1021 ATTGTTCATGTCGATGCTCTCGCTGTCATCATAGGCTCCCGGACGCGTGGCGCTTTC 1080
Db 1021 ATTGTTCATGTCGATGCTCTCGCTGTCATCATAGGCTCCCGGACGCGTGGCGCTTTC 1080
```

```
Db 2814565 ATTGTTCATGTCGATGCTCTCGCTGTCATCATAGGATCCCGGACGCGTGGCGCTTTC 2814506
QY 1081 TTGCTGCGGACGTCGCCGGAACGATATGTTGATGTATGCGGACGCGCTGCAATTTGG 1140
Db 2814505 TTGCTGCGGACGTCGCCGGAACGATATGTTGATGTATGCGGACGCGCTGCAATTTGG 2814446
QY 1141 CCGATGGAACGCTGATCTAGTCTGTTCGCAACGAGCTTTTTCATCAACGACGCGCATAG 1200
Db 2814445 CCGATGGAACGCTGATCTAGTCTGTTCGCAACGAGCTTTTTCATCAACGACGCGCATAG 2814386
QY 1201 AAGCCGCGGAATGCGACAGCGTGAATCAATTTGGGCGTTTGAACGTCACCGTCAACG 1260
Db 2814385 AAGCCGCGGAATGCGACAGCGTGAATCAATTTGGGCGTTTGAACGTCACCGTCAACG 2814326
QY 1261 ATGAGTGGCCCGTCTGGCGATGCGGCGCGCCATTCGCGGGGCAACGACACTAGCGA 1320
Db 2814325 ATGAGTGGCCCGTCTGGCGATGCGGCGCGCCATTCGCGGGGTCACGACACTAGCGA 2814266
QY 1321 GTCAATTGTGACACAGCGGCGGAGGATGAGCGCTTATGAGCGGACCGTCCACACT 1380
Db 2814265 GTCAATTGTGACACAGCGGCGGAGGATGAGCGCTTATGAGCGGACCGTCCACACT 2814206
QY 1381 CAACCGTCGCGCTTAAATTCCTCGCTCTTGGCGATCAGGAATCGCTAAATCCACCGCT 1440
Db 2814205 CAACCGTCGCGCTTAAATTCCTCGGACTTAGGATCAGGAATGCTTAAATCCACCGCT 2814146
QY 1441 CAAATTTTGTGTCTTTGATCTCAACCGGCGAGGTGCAAAAGGTCCATTTAGTATCAAG 1500
Db 2814145 TAAATTTTGTGTCTTTGATCTCAACCGGCGAGGTGCAAAAGGTCCATTTAGTATCAAG 2814086
QY 1501 TACTTAAATGTCAGTAAATCATCTGAT--TTTAAAGAGTCCACCATGAAATCACT 1559
Db 2814085 TAAATTTTGTGTCTTTGATCTCAACCGGCGAGGTGCAAAAGGTCCATTTAGTATCAAG 2814026
QY 1559 ATCTGCAAAAGCAGCAAGAAAGTGGCAAGAGCTTGAAGTCTTAATGCAACCTTCGCC 1619
Db 2814025 ATCTGCAAAAGCAGCAAGAAAGTGGCAAGAGCTTGAAGTCTTAATGCAAGTCCATGCA 2813966
QY 1619 AACAAAGGTGGAACCTTGGGCTTGGCAACAGATCTCTCAACAGTGAATCTTACCAAG 1679
Db 2813965 AACAAAGGTGGAACCTTGGGCTTGGCAACAGATCTCTCAACAGTGAATCTTACCAAG 2813906
QY 1679 CTATTCGATGATGAAGTGGGAAAGTGTCAATTCAGAACTGCAAGGCAATTCCTTGTG 1739
Db 2813905 CTATTCGATGATGAAGTGGGAAAGTGTCAATTCAGAACTGCAAGGCAATTCCTTGTG 2813846
QY 1739 GATGAATACGTGGGATTAACCGCGACGATGAAAAAGCTTCAAAAACATTCGTAA 2813786
Db 2813845 GATGAATACGTGGGATTAACCGCGACGATGAAAAAGCTTCAAAAACATTCGTAA 2813726
QY 1800 GAGTTCACTGACCATGCGATGATGTTGATGAAGAGTCTACAGCCGAGTGGTGAAC 1859
Db 2813725 GAGTTCACTGACCATGCGATGATGTTGATGAAGAGTCTACAGCCGAGTGGTGAAC 2813666
QY 1859 CCGATGATCAAGAAAGCTGCAAGATGAGGCAAAAGATCGCTGAGATTCGTTGAA 1919
Db 2813665 CCGATGATCAAGAAAGCTGCAAGATGAGGCAAAAGATCGCTGAGATTCGTTGAT 2813606
QY 1919 GTTCAATCTCTTGGATCGGCGGAAACGG--CACATCGCTTCAATTAACATCATCTTCT 1978
Db 2813605 GTTCAATCTCTTGGATCGGCGGAAACGG--CACATCGCTTCAATTAACATCATCTTCT 2813546
QY 1978 CTGTCAAGATGCAAAAGTTCAGGCGGTGACCTTAAACCTGTGAGAGCAAGGCTCGA 2038
Db 2813545 CTGTCAAGATGCAAAAGTTCAGGCGGTGACCTTAAACCTGTGAGAGCAAGGCTCGA 2813486
QY 2038 TTCTTCAACACCATGGAAGAGTCCCAACCCACCGCTCAACGAGGTTTGGGCACTTTG 2098
Db 2813485 TTCTTCAACACCATGGAAGAGTCCCAACCCACCGCTCAACGAGGTTTGGGCACTTTG 2813426
QY 2098 TCCGCGCGCAAAACATCTGTGTGTGGCAACTGTGTAAGGAAAGCGCATCCGC 2158
Db 2813425 TCCGCGCGCAAAACATCTGTGTGTGGCAACTGTGTAAGGAAAGCGCATCCGC 2813366
```

QY 2159 GGAATCTGTGGAAAGCCCACTGACTGCTTCTTGTGCCAGGTTCCATCTCTGTAGATGCAAC 2218
 DB 2813425 GGAATCTGTGGAAAGCCCACTGACTGACTGCTTCTTGTGCCAGGTTCCATCTCTGTAGATGCAAC 2813366
 QY 2219 ATGCCACATCACTGTTGGATGAAGCAGAGTATCCAGCTGGAAAACGCTGATCACTAC 2278
 DB 2813365 AATGCCACATCACTGTTGGATGAAGCAGAGTATCCAGCTGGAAAACGCTGATCACTAC 2813306
 QY 2279 GGTCTCATGAGAGCAATTTAAAGCTGCGCTAGAAAACAAAAGAAAGTACTGTGTGGGCTA 2338
 DB 2813305 CCTCTCATGAGAGCAATTTAAAGCTGCGCTAGAAAACAAAAGAAAGTACTGTGTGGGCTA 2813246
 QY 2339 TGCACACAGAACTTTCCAGTTTGGCCCTGCGCTACATGTGACTCTCTCCGACGGGAGAGC 2398
 DB 2813245 TGCACACAGAACTTTCCAGTTTGGCCCTGCGCTACATGTGACTCTCTCCGACGGGAGAGC 2813186
 QY 2399 TCAATGATCCCAACGGAAATGTAGCTGATGAGAGATACCTCCACGCTCTACTACACAGCAG 2458
 DB 2813185 TCAATGATCCCAACGGAAATGTAGCTGATGAGAGATACCTCCACGCTCTACTACACAGCAG 2813126
 QY 2459 ATCCAGGTTTCCCTTGTGACCAACAGCGCAACCGGCTGGCTCAACACACACACCGCGTTGA 2518
 DB 2813125 ATCCAGGTTTCCCTTGTGACCAACAGCGCAACCGGCTGGCTCAACACACACACCGCGTTGA 2813066
 QY 2519 CCGGACCGGACGAGTTGCGAGTGAAGCGCACCTGCGGACGCTCTTTACCCGGATGACATCTCT 2578
 DB 2813065 CCGGACCGGACGAGTTGCGAGTGAAGCGCACCTGCGGACGCTCTTTACCCGGATGACATCTCT 2813006
 QY 2579 ATGACCTGATGAGATGCTATTCGCGGTGAGCGGATTTTATCTGACGCGACACTTAACTTT 2638
 DB 2813005 ATGACCTGATGAGATGCTATTCGCGGTGAGCGGATTTTATCTGACGCGACACTTAACTTT 2812946
 QY 2639 TCTACACCGGCAACTTAAATTTGACGAAAGCGCGCGCAACCCAAAACCTTGTGAAAG 2698
 DB 2812945 TCTACACCGGCAACTTAAATTTGACGAAAGCGCGCGCAACCCAAAACCTTGTGAAAG 2812886
 QY 2699 TCGAGACCCCACTGGGCTGATGAGGCGGCAATTCATGCGGCTTCAAAAATCCGCTTA 2758
 DB 2812885 TCGAGACCCCACTGGGCTGATGAGGCGGCAATTCATGCGGCTTCAAAAATCCGCTTA 2812826
 QY 2759 TCGACGGAACCGGCGACGCGGTTTCAACCCCATTAACCGGATCCCATGATCAAGCCCTGATG 2818
 DB 2812825 TCGACGGAACCGGCGACGCGGTTTCAACCCCATTAACCGGATCCCATGATCAAGCCCTGATG 2812766
 QY 2819 GTGATGTTGGAACATGTTCTTGGGCGCAACCGCAAAAACCTCAACCGGTGACGCGGTTG 2878
 DB 2812765 GTGATGTTGGAACATGTTCTTGGGCGCAACCGCAAAAACCTCAACCGGTGACGCGGTTG 2812706
 QY 2879 TATACCGCTCGACAGATCTTGAATACTGGGAATTTCTCGGTGAAATCACTTTGACCTCA 2938
 DB 2812705 TATACCGCTCGACAGATCTTGAATACTGGGAATTTCTCGGTGAAATCACTTTGACCTCA 2812646
 QY 2939 GTGATGACAAACCTGTTCTGCTCTGATCTGTTCCGATGGCTCAATGTGGGAATGTC 2998
 DB 2812645 GTGATGACAAACCTGTTCTGCTCTGATCTGTTCCGATGGCTCAATGTGGGAATGTC 2812586
 QY 2999 CCAACCTTTTACGCTTCCGATGAGAAACCTGCGAAGTCTCGACGTTGATTTTCT 3058
 DB 2812585 CCAACCTTTTACGCTTCCGATGAGAAACCTGCGAAGTCTCGACGTTGATTTTCT 2812526
 QY 3059 GTTCACAGAGATTTGACCCGATTCACAGATGAGTTTCTCACTACGCAAGCTCTGACCAAGT 3118
 DB 2812525 GTTCACAGAGATTTGACCCGATTCACAGATGAGTTTCTCACTACGCAAGCTCTGACCAAGT 2812466
 QY 3119 GCGGATATGTCGTGCAAGGCTTGAAGAAACGACTTTCGCGCTTGGAGAGATTGAGG 3178
 DB 2812465 GCGGATATGTCGTGCAAGGCTTGAAGAAACGACTTTCGCGCTTGGAGAGATTGAGG 2812406
 QY 3179 AGCTGGAATTTCCGCAATGATTTCTACGCAACCGCAGGTTGCAAGTAAACGTTCTGATGCT 3238
 DB 2812405 AGCTGGAATTTCCGCAATGATTTCTACGCAACCGCAGGTTGCAAGTAAACGTTCTGATGCT 2812346

QY 3239 GGTCTGTGGGCTGATGAGGCGCTGCCGCGACAGATGATACCCAAACAGTTGACAGAGAG 3298
 DB 2812345 GGTCTGTGGGCTGATGAGGCGCTGCCGCGACAGATGATACCCAAACAGTTGACAGAGAG 2812286
 QY 3299 GATGGGTGACCTGCTGACTGTGCCCGCAAGCTTCAATTTGGCAACCAACGCGATCTAC 3358
 DB 2812285 GATGGGTGACCTGCTGACTGTGCCCGCAAGCTTCAATTTGGCAACCAACGCGATCTAC 2812226
 QY 3359 AAGACTCTCTTCCCAAGGCGGAGTGGGGGTATCAGATCTGTATTAGTTCTGAAC 3418
 DB 2812225 AAGACTCTCTTCCCAAGGCGGAGTGGGGGTATCAGATCTGTATTAGTTCTGAAC 2812166
 QY 3419 CTGTCCGATGACATCCGAGGCAATTTTCCCTGAGGAGATGAGTGCCTGTTGCTG 3478
 DB 2812165 CTGTCCGATGACATCCGAGGCAATTTTCCCTGAGGAGATGAGTGCCTGTTGCTG 2812106
 QY 3479 TGGATCTGATGATGATGCTGCGCTGATGAGTAAACCTGCGCAATTAAGTATGATCGCG 3538
 DB 2812105 TGGATCTGATGATGATGCTGCGCTGATGAGTAAACCTGCGCAATTAAGTATGATCGCG 2812046
 QY 3539 AGGATATACAGCCATTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3598
 DB 2812045 AGGATATACAGCCATTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2811987
 QY 3599 GCTTCAAGGTCAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 3658
 DB 2811986 GCTTCAAGGTCAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 2811927
 QY 3659 GTACAAATTAATTCGCAAAATCCCTGATGCGACACAAATTAACAGGTTTAAATTTGTTA 3718
 DB 2811926 GTACAAATTAATTCGCAAAATCCCTGATGCGACACAAATTAACAGGTTTAAATTTGTTA 2811867
 QY 3719 GCTTTTGAACAATTCATGTCGATATTTTGTCTTCCGCTTAAAGAGAAATTC 3778
 DB 2811866 GCTTTTGAACAATTCATGTCGATATTTTGTCTTCCGCTTAAAGAGAAATTC 2811807
 QY 3779 ATGACCAATTAAGACCTGCGGACGATCCGCGGACATTTGGGCGGAGACAACTT 3838
 DB 2811806 ATGACCAATTAAGACCTGCGGACGATCCGCGGACATTTGGGCGGAGACAACTT 2811747
 QY 3839 GTGCGCGCGGACACATGTCGCAACGCGTTTACGCTGTCGCTCAAAACCAAGAGATGTC 3898
 DB 2811746 GTGCGCGCGGACACATGTCGCAACGCGTTTACGCTGTCGCTCAAAACCAAGAGATGTC 2811687
 QY 3899 GATCGCAAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3958
 DB 2811686 GATCGCAAAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2811627
 QY 3959 TTTCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4018
 DB 2811626 TTTCCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2811567
 QY 4019 ACCCTCAAAAGACATGCTGTGTCACACAGACGCTCAAAAGATGTTGGCTAAACGCG 4078
 DB 2811566 ACCCTCAAAAGACATGCTGTGTCACACAGACGCTCAAAAGATGTTGGCTAAACGCG 2811507
 QY 4079 AACTGATGACGCTGTGTGAAGGATTTGCGGACATTTTCTCCCGCTGATTCATATC 4138
 DB 2811506 AACTGATGACGCTGTGTGAAGGATTTGCGGACATTTTCTCCCGCTGATTCATATC 2811447
 QY 4139 TTTGGTTGGGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4198
 DB 2811446 TTTGGTTGGGGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2811387
 QY 4199 GGTCCGCAATCACTGATGAGATGTTTCCCTGATGATGATGATGATGATGATGATGATGATG 4258
 DB 2811386 GGTCCGCAATCACTGATGAGATGTTTCCCTGATGATGATGATGATGATGATGATGATGATG 2811327
 QY 4259 CTGATGACATCTGCGCGCTTCTGCTTCTGCAAGTGTGTTGTTTCAACGCAACAG 4318
 DB 2811326 CTGATGACATCTGCGCGCTTCTGCTTCTGCAAGTGTGTTGTTTCAACGCAACAG 2811267
 QY 4319 CGTTTCGATGAGATTAAGATTTCTGGGCGCGGCAATTTGATGATGATGATGATGATGATG 4378

```

Db      2811266 CGTTTCGGGGGCAATGAGTTCTCGGGGCGCGGATTTGGATGATCGATGTTTCCGAGC 2811207
Qy      4379 CTGGTTAACGGCTTACGACGTCGCCCGCACCATGACCGCGGCGGAAATGTCATGTC 4438
Db      2811206 TTGGTAACGGCTTACGACGTCGCCCGCACCATGACCGCGGCGGAAATGTCATGTC 2811147
Qy      4439 CTGGTTTGGATGATGTTGCTCAAGCTGGTTACAGGGGACGCTGCTTCTGTCGTCGG 4498
Db      2811146 CTGGTTTGGATGATGTTGCTCAAGCTGGTTACAGGGGACGCTGCTTCTGTCGTCGG 2811087
Qy      4499 GTCTCTGGATTCTGGCAACGATCGAAGATTCTGCAACAGCGACTGAGGCACTGCA 4558
Db      2811086 GTTTCCTGGATTCTGGCAACGATCGAAGATTCTGCAACAGCGACTGAGGCACTGCA 2811027
Qy      4559 GACTTCCTGATCAACCCAGTGTGACTCTGCTGCTCAACGGCTTCTTTCGTTCAATGCT 4618
Db      2811026 GACTTCCTGATCACTCCAGTGTGACTCTGCTGCTCAACGGATTCCTTACATTCG 2810967
Qy      4619 ATTGGTCCAGCAATGCGCTGGGATGGGATGACTTGTGGCAACAGGCTGCGAGGACCTAT 4678
Db      2810966 ATTGGCCCAAGCAATGCGCTGGGATGGGATGACTTGTGGCAACAGGCTGCGAGGAC 2810907
Qy      4679 GATTTCGATGATCGAGTCGCGGATGCTGCTTTCGCTGCTGCTCAACCAATCGTTATC 4738
Db      2810906 GATTTCGATGATCGAGTCGCGGATGCTGCTTTCGCTGCTGCTCAACCAATCGTTATC 2810847
Qy      4739 ACTGGTCTGACCAATGCTTTCGCGCAATGAGCTGAGCTGTTTCAACAGGCTGATCC 4798
Db      2810846 ACTGGTCTGACCAATGCTTTCGCGCAATGAGCTGAGCTGTTTAAACAGGCTGATCC 2810787
Qy      4799 TTCACTTTCGCAACCGCATCTGATGCGCAATATCGCGACAGGATGTTTGGCAAGT 4858
Db      2810786 TTCACTTTCGCAACCGCATCTGATGCGCAATATCGCGACAGGATGTTTGGCAAGT 2810727
Qy      4859 TTCTTCTAGCGAAGTGAAGTCAAGGACCTTTCGAGGCTTCCTCCGCT 4918
Db      2810726 TTCTTCTAGCGAAGTGAAGTCAAGGACCTTTCGAGGCTTCCTCCGCT 2810667
Qy      4919 GTTCTTGGATTAACAGAGCTGCGATCTTCGCTGTAAGACCTTCGCTGCGCTGCGCTTC 4978
Db      2810666 GTTCTTGGATTAACAGAGCTGCGATCTTCGCTGTAAGACCTTCGCTGCGCTGCGCTTC 2810607
Qy      4979 TACATTGGATGATGATGACGACGATATCGGTGGGCTTTCATTTGATGATCAAG 5038
Db      2810606 TACATTGGATGATGATGACGACGATATCGGTGGGCTTTCATTTGATGATCAAG 2810547
Qy      5039 GCAAGTTCGCTGGGCGCTGCAAGTTCCTTGGGCTTTCATTTGATGATCAAG 5098
Db      2810546 GCAAGTTCGCTGGGCGCTGCAAGTTCCTTGGGCTTTCATTTGATGATCAAG 2810487
Qy      5099 GTCAATGTTTGGTGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 5158
Db      2810486 GTCAATGTTTGGTGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 2810427
Qy      5159 TATGGCTTACTTGGTGGCGCAACGCGACATGATCGATGCAACCGCTGCTCA 5218
Db      2810426 TATGGCTTACTTGGTGGCGCAACGCGACATGATCGATGCAACCGCTGCTCA 2810367
Qy      5219 GTGCTCGAGGAACGACCAAGCCGAGAGGACCGGCAAGATTTTCAACGATTC 5278
Db      2810366 GTGCTCGAGGAACGACCAAGCCGAGAGGACCGGCAAGATTTTCAACGATTC 2810307
Qy      5279 ACCATCATTCAGAGCACTTTGACCGGTGAAGCTATGCACTGAGAGGCTGACGATGCG 5338
Db      2810306 ACCATCATTCAGAGCACTTTGACCGGTGAAGCTATGCACTGAGAGGCTGACGATGCG 2810247
Qy      5339 ATGTTTGCAGGCGGAAAGCTTGGCTCAGGCTTGGGATGATGATGATGATGATGATGATGAT 5398
Db      2810246 ATGTTTGCAGGCGGAAAGCTTGGCTCAGGCTTGGGATGATGATGATGATGATGATGATGAT 2810187
Qy      5399 GTTTCACAGTGAAGGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 5458

```

```

Db      2810186 GTTTCACGAGTGAAGGAAAGATGATGATGATGATGATGATGATGATGATGATGAT 2810127
Qy      5459 CGCATTAAGGCTGAGAGATGTTTCCATATGATATCTTGAATGACATGTTTGCAGACC 5518
Db      2810126 CGCATTAAGGCTGAGAGATGTTTCCATATGATATCTTGAATGACATGTTTGCAGACC 2810067
Qy      5519 GTAAACCTCAACGAGGACGATTTAACCGCTGGAAGAGAGGAGGATGATGATGATGATGATGAT 5578
Db      2810066 GTAAACCTCAACGAGGACGATTTAACCGCTGGAAGAGAGGAGGATGATGATGATGAT 2810007
Qy      5579 GGGAGCTGCTGTGTAATTCGATATGATGATGATGATGATGATGATGATGATGATGATGAT 5638
Db      2810066 GGGAGCTGCTGTGTAATTCGATATGATGATGATGATGATGATGATGATGATGATGAT 2809947
Qy      5639 ACGCGATGTTGTTTTCGAAATTAAGAAACCGGACCTGTAACACTTACGTTTGGGC 5698
Db      2809946 ACGCGATGTTGTTTTCGAAATTAAGAAACCGGACCTGTAACACTTACGTTTGGGC 2809887
Qy      5699 GAAATTTGAAGCGGAGCAACCTGCTCAACGTCGCAAGAAAGAAAGCGGTCAGCAACA 5758
Db      2809886 GAAATTTGAAGCGGAGCAACCTGCTCAACGTCGCAAGAAAGAAAGCGGTCAGCAACA 2809827
Qy      5759 CCATTAAGTTGAAACCTTGAATGTTGCAACACAGTTAGCTAGGGAGCTGATCTTACGC 5818
Db      2809826 CCATTAAGTTGAAACCTTGAATGTTGCAACACAGTTAGCTAGGGAGCTGATCTTACGC 2809767
Qy      5819 ATCTTGAACACCGGATACCGGATACGTTGATGATGATGATGATGATGATGATGATGATGATGAT 5878
Db      2809766 ATCTTGAACACCGGATACCGGATACGTTGATGATGATGATGATGATGATGATGATGAT 2809707
Qy      5879 TCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5938
Db      2809706 TCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2809647
Qy      5939 GCAAGTACGTTGATATTTTGGCGCGCTG 5967
Db      2809646 GCAAGTACGTTGATATTTTGGCGCGCTG 2809618

```

```

RESULT 2
US-09-738-626-2904
/ Sequence 2904, Application US/09738626
/ Publication No. US20020197605A1
GENERAL INFORMATION:
/ APPLICANT: NAKAGAWA, SATOSHI
/ APPLICANT: MIZOGUCHI, HIROSHI
/ APPLICANT: ANDO, SEIRO
/ APPLICANT: HAYASHI, MIKIRO
/ APPLICANT: OCHIAI, KEIKO
/ APPLICANT: YOKOI, HARUHIKO
/ APPLICANT: TATEISHI, NAOKO
/ APPLICANT: SENOH, AKIHIRO
/ APPLICANT: IKEDA, MASATO
/ APPLICANT: OZAKI, AKIO
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-125
/ CURRENT APPLICATION NUMBER: US/09/738,626
/ CURRENT FILING DATE: 2000-12-18
/ PRIOR APPLICATION NUMBER: JP 99/377484
/ PRIOR FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: JP 00/159162
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: JP 00/280988
/ PRIOR FILING DATE: 2000-08-03
/ NUMBER OF SEQ ID NOS: 7059
/ SOFTWARE: PatentIn ver. 3.0
/ SEQ ID NO 2904
/ LENGTH: 1983
/ TYPE: DNA
/ ORGANISM: Corynebacterium glutamicum
US-09-738-626-2904

```

Query Match 31.4%; Score 1874.2; DB 9; Length 1983;

Best Local Similarity 96.6%; Pred. No. 0;
Matches 1915; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```

OY 3779 ATGGACCATTAAGAGACCTCGCGCAACCGATCTCGCGACATTTGGCGCGCGCAAGCAACATT 3838
DB 1 ATGGACCATTAAGAGACCTCGCGCAACCGATCTCGCGACATTTGGCGCGCGCAAGCAACATT 60
OY 3839 GTGCGCGCGCACACTGTGCAACGCGCTTACGCTCGTGTCTCAAGACACCAAGATGTG 3898
DB 61 GTGCGCGCGCACACTGTGCAACGCGCTTACGCTCGTGTCTCAAGACACCAAGATGTG 120
OY 3899 GATCGCAAAAGTCTGATGATGATCCAGATCTGAAAAGCACCTTTGAAACTGCGCGCATG 3958
DB 121 GATCGCAAAAGTCTGATGATGATCCAGATCTGAAAAGCACCTTTGAAACTGCGCGCATG 180
OY 3959 TTCCAGATATCGTCGCGCGCGAGCGAGTGGAGATCATGTTTCAAAAGACTCGATGACGA 4018
DB 181 TTCCAGATATCGTCGCGCGCGAGCGAGTGGAGATCATGTTTCAAAAGACTCGATGACGA 240
OY 4019 ACCTCAAAAGACATCGCTGTGTCCACAGAGACGCTCAAAAGATGTTGGCTAAACAGCGC 4078
DB 241 ACCTCAAAAGACATCGCTGTGTGTCCACAGAGACGCTCAAAAGATGTTGGCTAAACAGCGC 300
OY 4079 AACTGTTCAAGCCGTGCTGTGAAGGTAATGGCGGACATTTTCGTCGCTGATTCACATC 4138
DB 301 AACTGTTCAAGCCGTGCTGTGAAGGTAATGGCGGACATTTTCGTCGCTGATTCACATC 360
OY 4139 TTGGTTGGTGGCGGTGTGTCTCATAGGCTATCAAAATGTGTGGTGGCGAGATCTGTTC 4198
DB 361 TTGGTTGGTGGCGGTGTGTCTCATAGGCTATCAAAATGTGTGGTGGCGAGATCTGTTC 420
OY 4199 GGTCCCAATCACTGTGTGAGATGTTCCCTCAGATCAGCGGTGTGTGCTGAGATGATCAAC 4258
DB 421 GGTCCCAATCACTGTGTGAGATGTTCCCTCAGATCAGCGGTGTGTGCTGAGATGATCAAC 480
OY 4259 CTGATGGCATCTGCGCGCTTCCGCTTCTTGCCAGATGTTGGTGGTTCACCGCAACCAAG 4318
DB 481 CTGATGGCATCTGCGCGCTTCCGCTTCTTGCCAGATGTTGGTGGTTCACCGCAACCAAG 540
OY 4319 CGTTTCGGTGGCAATAGATTTCCCTGCGCGCGCGGATGTTGACATGCTGTTCCCAAC 4378
DB 541 CGTTTCGGTGGCAATAGATTTCCCTGCGCGCGCGGATGTTGACATGCTGTTCCCAAC 600
OY 4379 CTGGTTAACGCGCTACGACGTGCGCGCACCATGACCGCGGCGCAATGTCATGAGTGC 4438
DB 601 CTGGTTAACGCGCTACGACGTGCGCGCACCATGAGCTGCGGCGCAATGTCATGAGTGC 660
OY 4439 CTGTTTGGTTGGATGTTGCTCAAGCTGTGTACACAGGCGACCGTGTCTCTGTGCTGTG 4498
DB 661 CTGTTTGGTTGGATGTTGCTCAAGCTGTGTACACAGGCGACCGTGTCTCTGTGCTGTG 720
OY 4499 GTCTCTGATGATCTGCGCAACGATCGAAGATTTCTGCAACAAGGACTCATGCGCACTGCA 4558
DB 721 GTCTCTGATGATCTGCGCAACGATCGAAGATTTCTGCAACAAGGACTCATGCGCACTGCA 780
OY 4559 GACTTCCTGATACCCGAGTGTGAATCTGCTGCTGACACCGGCTTCTTACGTTCAATGCT 4618
DB 781 GACTTCCTGATACCCGAGTGTGAATCTGCTGCTGACACCGGCTTCTTACGTTCAATGCT 840
OY 4619 ATTGTCAGCAATGCGCTGTGTGAGCTTGTGCTGCAACAGGCTGTCAGAGGACTCTAT 4678
DB 841 ATTGTCAGCAATGCGCTGTGTGAGCTTGTGCTGCAACAGGCTGTCAGAGGACTCTAT 900
OY 4679 GATTTTCGGTGTGCAATGCGCGGTCTGTCTTTTCGCTGTGCTACTACCAATCGTTATC 4738
DB 901 GATTTTCGGTGTGCAATGCGCGGTCTGTCTTTTCGCTGTGCTACTACCAATCGTTATC 960
OY 4739 ACTGCTCTGACCAAGTCTTCCGCGCAATGAGCTGAGCTGTTCACCAAGGCTGGAATCC 4798
DB 961 ACTGCTCTGACCAAGTCTTCCGCGCAATGAGCTGAGCTGTTCACCAAGGCTGGAATCC 1020
OY 4799 TTTCATCTTGGCAACGCGATCATTGCGCAATATGCGCGAGGCTGACAGATGTTTGGCAGTG 4858

```

```

DB 1021 TTTCATCTTGGCAACGCGATCTATGCTAATATGCGCCAGGCTGCGCATGTTTGGCAGTG 1080
OY 4859 TTCTTCCTTACGAAAGGTGTAAGTCAAGGCGCTTGCAGGCTTCAAGTGTCTCCGCT 4918
DB 1081 TTCTTCCTTACGAAAGGTGTAAGTCAAGGCGCTTGCAGGCTTCAAGTGTCTCCGCT 1140
OY 4919 GTTCTTGTATTTACAGAGCTCGATCTCTCGGTGTGAACCTTTCGCTGCGTGGCGCTTC 4978
DB 1141 GTTCTTGTATTTACAGAGCTCGATCTCTCGGTGTGAACCTTTCGCTGCGTGGCGCTTC 1200
OY 4979 TACATTTGATCGGTACCGGACCTATCGGTGCGCTTGTGATGCACTCTTGTATATCAAG 5038
DB 1201 TACATTTGATCGGTACCGGACCTATCGGTGCGCTTGTGATGCACTCTTGTATATCAAG 1260
OY 5039 GCAATTTGCTTGGCGCTGCAAGTCTTCTGCTGTGTTTCTTATGATGCTCAATATG 5098
DB 1261 GCAATTTGCTTGGCGCTGCAAGTCTTCTGCTGTGTTTCTTATGATGCTCAATATG 1320
OY 5099 GTCATGTTCTTGGTTTGCAGGTAAGTACCTTGTGATCGCATTCGCGCGACGATTTGCT 5158
DB 1321 GTCATGTTCTTGGTTTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGT 1380
OY 5159 TATGCTTCTTACTTGTGTGCGGCAACGCGACATTTGATCGATGCAACGCTGTCTCA 5218
DB 1381 TATGCTTCTTACTTGTGTGCGGCAACGCGACATTTGATCGATGCAACGCTGTCTCA 1440
OY 5219 GTGCTTGCAGGAACGACCAAGCGAAGGAGAACCGCGGAAATTTTCAACGATTC 5278
DB 1441 GTGCTTGCAGGAACGACCAAGCGAAGGAGAACCGCGGAAATTTTCAACGATTC 1500
OY 5279 ACCATCATCAGGACCTTTGACCGGTGAAGTATCGACTGAGCGCTGACGATTC 5338
DB 1501 ACCATCATCAGGACCTTTGACCGGTGAAGTATCGACTGAGCGCTGACGATTC 1560
OY 5339 ATGTTTGCAGGGAAGGCTTGTGCTCAAGTGTGTGATGCTGCCACCAAGGCGGAGCTG 5398
DB 1561 ATGTTTGCAGGGAAGGCTTGTGCTCAAGTGTGTGATGCTGCCACCAAGGCGGAGCTG 1620
OY 5399 GTTTCACCAAGTGAAGGGAAGATCGTGTGCTTCCATCTGTGTCAACGCTTTCGAGTC 5458
DB 1621 GTTTCACCAAGTGAAGGGAAGATCGTGTGCTTCCATCTGTGTCAACGCTTTCGAGTC 1680
OY 5459 CGCATTGAAGGCTGAAGATGTTGCAATGATGATCTTGTGATGATGATGATGATGATG 5518
DB 1681 CGCATTGAAGGCTGAAGATGTTGCAATGATGATGATGATGATGATGATGATGATGATG 1740
OY 5519 GTTAACTTCAACGCGACGACTTTAACCGGCTGAAGGAAGCAGGCGGATGTAAGTCAAGCA 5578
DB 1741 GTTAACTTCAACGCGACGACTTTAACCGGCTGAAGGAAGCAGGCGGATGTAAGTCAAGCA 1800
OY 5579 GGGAGGCTGCTGTGTAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 5638
DB 1801 GGGAGGCTGCTGTGTAATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1860
OY 5639 ACGCGATGTTGTTTGTGCAATTAACAAGAAACCGGACCTGTAACATTTAGGTTTGGGC 5698
DB 1861 ACGCGATGTTGTTTGTGCAATTAACAAGAAACCGGACCTGTAACATTTAGGTTTGGGC 1920
OY 5699 GAAATTTGAAGCGGAGCAACCTGCTCAACGTCGCAAGAAAGAAAGAGCGGTGCAACCA 5758
DB 1921 GAAATTTGAAGCGGAGCAACCTGCTCAACGTCGCAAGAAAGAAAGAGCGGTGCAACCA 1980
OY 5759 CCA 5761
DB 1981 CCA 1983

```

RESULT 3
US-10-450-055-9
; Sequence 9, Application US/10450055
; Publication No. US20040043953A1
; GENERAL INFORMATION:
; APPLICANT: BASF Aktiengesellschaft

;; TITLE OF INVENTION: No. US20040043953A1e1 genes of Corynebacterium
;; FILE REFERENCE: 936, 2000
;; CURRENT APPLICATION NUMBER: US/10/450,055
;; CURRENT FILING DATE: 2003-06-10
;; NUMBER OF SEQ ID NOS: 46
;; SOFTWARE: PatentIn Vers. 2.0
;; SEQ ID NO 9
;; LENGTH: 1527
;; TYPE: DNA
;; ORGANISM: Corynebacterium glutamicum
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (101)..(1504)
;; OTHER INFORMATION: KAS00315
;; US-10-450-055-9

Query Match 23.7%; Score 1412.6; DB 17; Length 1527;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1462; Conservative 0; Mismatches 64; Indels 1; Gaps 1;

QY 4259 CTGATGCAATGCGCCGTTCCGTTCTTGCAGTGTGTTGTTTCCGCCAACAG 4318
DB 1 CTATGCAATGCGCCGTTCCGTTCTTGCAGTGTGTTGTTTCCGCCAACAG 60
QY 4319 CGTTTCGGTGGCAATGATTCCTGGGGCGCGGATGTGATGCGATGTGTTCCGAG 4377
DB 61 CGTTTCGGGCGCAATGATTCCTGGGGCGCGGATGTGATGCGATGTGTTCCGAG 120
QY 4378 CCGTGTAAACGGCTACGACGTGGCCGCCACATGACCGGGCGGAATGCCAATGTGTC 4437
DB 121 CTGTGGAAACGGCTACGACGTGGCCGCCACATGCGTGGGGGAAATGCCAATGTGTC 180
QY 4438 CCTGTGTTGTTGATGTGTTCTCAAGCTGTGTTCAAGGGACCGTGTCTGTGTGT 4497
DB 181 CCGTGTGTTGTTGATGTGTTCTCAAGCTGTGTTCAAGGGACCGTGTCTGTGTGT 240
QY 4498 GGTCTCTTGATTCGGGCAACATGAGAAATTCCTGCAACAGCGCATGATGGGCACTGC 4557
DB 241 GGTCTCTTGATTCGGGCAACATGAGAAATTCCTGCAACAGCGCATGATGGGCACTGC 300
QY 4558 AGACTTCTGATCAACCCGATGTGATCTGCTGCTCACCGCTTCCTTACGTTCACTTCG 4617
DB 301 AGACTTCTGATCAACCCGATGTGATCTGCTGCTCACCGATTCCTTACATTCATGCG 360
QY 4618 TATGTGTCAGCAATGCGTGGTGGTGACTTGTGGGCAACAGCTGTGAGGACTCTA 4677
DB 361 CATGGCCAGCAATGCGTGGTGGTGAATGTGTGGGCAACAGCTGTGAGGACTCTTA 420
QY 4678 TGATTTGGGAGTCAGTGGGCGGTCTGCTTTTGGTGTGATCTGACCAATCGTAT 4737
DB 421 TGATTTGGGAGTCAGTGGGCGGTCTGCTTTTGGTGTGATCTGACCAATCGTAT 480
QY 4738 CACTGCTGCAACCAAGTCTTCCCGCAATGAGCTGAGCTGTTCAACAGGATGATC 4797
DB 481 CACTGCTGCAACCAAGTCTTCCCGCAATGAGCTGAGCTGTTCAACAGGATGATC 540
QY 4798 CTTCATCTTGGCAACCGCATTCATGCGCAATATCGGCGAGGGTGCAGATGTTGGCAAT 4857
DB 541 CTTCATCTTGGCAACCGCATTCATGCGCAATATCGGCGAGGGTGCAGATGTTGGCAAT 600
QY 4858 GTTCTTCCAGCAAGAGTGAAGTCAAGGGCTTGGAGGCTTCAAGGTCTCCCGC 4917
DB 601 GTTCTTCCAGCAAGAGTGAAGTCAAGGGCTTGGAGGCTTCAAGGTCTCCCGC 660
QY 4918 TGTCTTGGTATTAACAAGCTGCGATCTTCCGTTGAACTTCCGCTGGCGCTGCGCT 4977
DB 661 TGTCTTGGTATTAACAAGCTGCGATCTTCCGTTGAACTTCCGCTGGCGCTGCGCT 720
QY 4978 CTTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5037
DB 721 CTTCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 5038 GGCAGTGGCTGGGGCGCTGCAAGTTCCTTGGGTGTTGTTTATGATGATGATGATGATGAT 5097

DB 781 GGCAGTGGCTGGGGCGCTGCAAGTTCCTTGGGTGTTGTTTATGATGATGATGATGATGATGAT 840
QY 5098 GGTCAATGTTCTTGGTGGCGGATAGTACCTTGTGATGATGATGATGATGATGATGATGATGAT 5157
DB 841 GGTCAATGTTCTTGGTGGCGGATAGTACCTTGTGATGATGATGATGATGATGATGATGATGAT 900
QY 5158 TTATGAGCTTTAATGTTGTTGCGCCGCAAGGAGATGATGATGATGATGATGATGATGATGATGAT 5217
DB 901 TTATGAGCTTTAATGTTGTTGCGCCGCAAGGAGATGATGATGATGATGATGATGATGATGATGAT 960
QY 5218 AGTGCCTGAGAAACGACCAAGCCGAGAGCAACCCGAGATTTCAAGATTC 5277
DB 961 AGTGCCTGAGAAACGACCAAGCCGAGAGCAACCCGAGATTTCAAGATTC 1020
QY 5278 CACCATCATCAGGACCTTTGACCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5337
DB 1021 CACCATCATCAGGACCTTTGACCGGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 5338 CATGTTTCCAGCGGAAGCTTGGCTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 5397
DB 1081 CATGTTTCCAGCGGAAGCTTGGCTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 5398 GGTTCACCAAGTGAACGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5457
DB 1141 AGTTTCTCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 5458 CCGCATGAGGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5517
DB 1201 CCGCATGAGGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
QY 5518 CGTAAACCTCAACCGGACGACCTTTAACCCGCTGAAGAGAGGCGATGATGATGATGATGATGATGAT 5577
DB 1261 CGTAAACCTCAACCGGACGACCTTTAACCCGCTGAAGAGAGGCGATGATGATGATGATGATGATGAT 1320
QY 5578 AGGGAGCTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5637
DB 1321 AGGGAGCTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 5638 CAGCGCATGTTGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5697
DB 1381 CAGCGCATGTTGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 5698 CGAATGAGGAGGAGCAACCTGCTCAAGTGCAGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 5757
DB 1441 CGAATGAGGAGGAGCAACCTGCTCAAGTGCAGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
QY 5758 ACCATGATGAAACCTTGAAGTTCG 5784
DB 1501 ACCATGATGAAACCTTGAAGTTCG 1527

RESULT 4
US-10-781-014-347
; Sequence 347, Application US/10781014
; Publication No. US20040180408A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Krieger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
; FILE REFERENCE: BGI-126CPN
; CURRENT FILING DATE: 2004-02-17
; PRIOR APPLICATION NUMBER: US 09/602,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/143,208

PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 60/151,572
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19931412.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931413.6
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931419.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931424.1
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931428.4
PRIOR FILING DATE: 1999-07-08
Remainder Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 784
SEQ ID NO 347
LENGTH: 1342
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: CDS
LOCATION: (101)..(1342)
OTHER INFORMATION: RXA00041
US-10-781-014-347

Query Match 22.1%; Score 1321.2; DB 18; Length 1342;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1329; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

2238 ATGAAGCAGCATGATTCAGCTGAGAAAGCGTGAATCACTACCGTCTCATGAGCAATTAA 2297
1 ATGAAGCAGCATGATTCAGCTGAGAAAGCGTGAATCACTACCGTCTCATGAGCAATTAA 60
2298 AGCTGCGCTAGAAACAAAGAAAGTAAGTGTGTGGGCTATGCAACAGAACTTTCCAG 2357
61 AGCTGCGCTAGAAACAAAGAAAGTAAGTGTGTGGGCTATGCAACAGAACTTTCCAG 120
2358 TTTGGCGCTGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2417
121 TTTGGCGCTGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
2418 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2477
181 GATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
2478 ACCAAGGCGCAGCGGCTGGGCTCACACACACACACACACACACACACACACACAC 2537
241 ACCAAGGCGCAGCGGCTGGGCTCACACACACACACACACACACACACACACACAC 300
2538 GTGAGCGCAGCTGCGCGGCTCTTTACCGGATGATGATGATGATGATGATGATGATGAT 2597
301 GTGAGCGCAGCTGCGCGGCTCTTTACCGGATGATGATGATGATGATGATGATGATGAT 360
2598 TTTCCGCTGAGCGCTATTTACTGAGCGCACATTAACTTTTCTAACCGGCAACTTAA 2657
361 TTTCCGCTGAGCGCTATTTACTGAGCGCACATTAACTTTTCTAACCGGCAACTTAA 420
2658 AATTGACGAAAGCGCGCGCCACCAAAACCTTGTGGAAGTGAAGACCAACCTGGGCT 2717
421 AATTGACGAAAGCGCGCGCCACCAAAACCTTGTGGAAGTGAAGACCAACCTGGGCT 480
2718 GATGGCGGCAATTCATGCGCTTCCGCTTAAATTCGCTTATGACGACCGCGCGCG 2777
481 GATGGCGGCAATTCATGCGCTTCCGCTTAAATTCGCTTATGACGACCGCGCGCG 540
2778 TTTTACACCCCATTTACCGCGATCCATGATCAGCGCTGATGATGATGATGATGATGAT 2837
541 TTTTACACCCCATTTACCGCGATCCATGATCAGCGCTGATGATGATGATGATGATGAT 600
2838 TCTGGGGCGCAACGCGAAACCTGACCGGTGACGGGTTCTTATACGCTCGACAGATCT 2897

601 TCTTGGGGCCCAACGCGAAACCTCACCGGTGACGGTGTATACCGCTCGACAGATCT 660
2898 TGAAGACCTGGAATTCCTCCGCTGAAATGACCTTTGACCTCATGATGACAACTGGTTC 2957
661 TGAAGACCTGGAATTCCTCCGCTGAAATGACCTTTGACCTCATGATGACAACTGGTTC 720
2958 TGTCTCTGATCTCGTCCCGATGAGTGTATGATGAGGAAATGCCCCAATCTTTTACGCTTCG 3017
721 TGTCTCTGATCTCGTCCCGATGAGTGTATGATGAGGAAATGCCCCAATCTTTTACGCTTCG 780
3018 CGATGAGAAACCTGCGAAGATCTTCGACGCTGTATTTCTGTCCACAGATTTGACCG 3077
781 CGATGAGAAACCTGCGAAGATCTTCGACGCTGTATTTCTGTCCACAGATTTGACCG 840
3078 AATCAGATGAGTGTATCTCATGACGACCTTCGACGAGTGTATTTGCTGTGAGAA 3137
841 AATCAGATGAGTGTATCTCATGACGACCTTCGACGAGTGTATTTGCTGTGAGAA 900
3138 GCTTGAAGAAACGACCTTCGCGGTCTTGCAGAGATTCAGCGAGTGTATTTGCGGCATGA 3197
901 GCTTGAAGAAACGACCTTCGCGGTCTTGCAGAGATTCAGCGAGTGTATTTGCGGCATGA 960
3198 ATTCTACGACCGCAGTGTGAGTAACGCTTGTATGCTGTGCTGTGAGTGTGATGAG 3257
961 ATTCTACGACCGCAGTGTGAGTAACGCTTGTATGCTGTGCTGTGAGTGTGATGAG 1020
3258 GTTCCCGCGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3317
1021 GTTCCCGCGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
3318 TGTGCGCGCAGCTTCAATTTGCGGACCAACGCGATCTACCAAGAGCTCTTCTCCAGA 3377
1081 TGTGCGCGCAGCTTCAATTTGCGGACCAACGCGATCTACCAAGAGCTCTTCTCCAGA 1140
3378 GGGGAGTGTGGGAGTAATCAGATCTGATTAAGTGTGAACTGTCCAGATGACATCCG 3437
1141 GGGGAGTGTGGGAGTAATCAGATCTGATTAAGTGTGAACTGTCCAGATGACATCCG 1200
3438 AGGCAATTTTCCCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3497
1201 AGGCAATTTTCCCTCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
3498 TCGGTAGCTGAGTAACCTGCGCAATTAAGTATTCGCGACGATTAATACAGCAATTGA 3557
1261 TCGGTAGCTGAGTAACCTGCGCAATTAAGTATTCGCGACGATTAATACAGCAATTGA 1320
3558 GATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3579
1321 GATTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1342

RESULT 5
US-09-738-626-2905
Sequence 2905, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAMA, SATOSHI
APPLICANT: ANDO, SEIKO
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162

FEATURE:
NAME/KEY: CDS
LOCATION: (101)..(1264)
OTHER INFORMATION: RXN00043
US-10-781-014-413

Query Match 20.7%; Score 1237.4; DB 18; Length 1287;
Best Local Similarity 97.6%; Pred. No. 0;
Matches 1256; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 242 AACAGAGGCTCAAGTCCGAGATTAATTAAGTACCTAGATCCGTAGACATAGACATCATAC 301
DB 1 AACAGAGGCTCAAGTCCGAGATTAATTAAGTACCTAGATCCGTAGACATAGACATCATAC 60

QY 302 GTCTTAGTCTGTGTGAGAGGAAACCAATTAACCTAGAGATGCGAGAGTGTGATTA 361
DB 61 GTCTTAGTCTGTGTGAGAGGAAACCAATTAACCTAGAGATGCGAGAGTGTGATTA 120

QY 362 TCAAGAAAATGACGTCAGAGTAAATAATTTGAGGAGAAATGTTCCCGCCCTCG 421
DB 121 TCAAGAAAATGACGTCAGAGTAAATAATTTGAGGAGAAATGTTACCCCAACG 180

QY 422 GGTGATGATGAGCTTTCTCCAACTCGAAAACGCGATCATCAAGAACTCTGAGAAC 481
DB 181 GGTGATGATGAGCTTTCTCCAACTCGAAAACGCGATCATCAAGAACTCTGAGAAC 240

QY 482 AGCAGCTTAAATGACGATTCACCCCGAATCCCGACGATTTCCCGGTTTATTA 541
DB 241 AGCAGCTTAAATGACGATTCACCCCGAATCCCGACGATTTCCCGGTTTATTA 300

QY 542 TCTTCAATATCAGGAGGAAAGGAGGCGGTTCTTACGAGGAAACGAGACGAGCGAG 601
DB 301 TCTTCAATATCAGGAGGAAAGGAGGCGGTTCTTACGAGGAAACGAGACGAGCGAG 360

QY 602 GAACACCGGCGAGTATCACCGCAAGATGCGACGACGATGTTGCAAGATGTTTC 661
DB 361 GAATGCGGCGAGTATCACCGCAAGATGCGACGACGATGTTGCAAGATGTTTC 420

QY 662 GCGCGCGGCTGACGACTGCGACGCGAGTGAAGAACTTATTCCTTGTGTGAAGGT 721
DB 421 GCGCGCGGCTGACGACTGCGACGCGAGTGAAGAACTTATTCCTTGTGTGAAGGT 480

QY 722 CCGCTGTGCGGATTCACCTGAGGCGCTTTCATCAACGATGCGGTGTGTCTCA 781
DB 481 CCGCTGTGCGGATTCACCTGAGGCGCTTTCATCAACGATGCGGTGTGTCTCA 540

QY 782 AAACCGGATTCATTTTCCCGGCAACCAAGATCTTCCCGGATTCATTCATGCTTC 841
DB 541 AAACCGGATTCATTTTCCCGGCAACCAAGATCTTCCCGGATTCATTCATGCTTC 600

QY 842 AAAGGTTGATCAATTCGATCAAGTAGCGCGGAACTGACATCTTTTGAAGCTTC 901
DB 601 AAAGGTTGATCAATTCGATCAAGTAGCGCGGAACTGACATCTTTTGAAGCTTC 660

QY 902 CGATCTCTGCGGACCGACACATCATTTGCTTCCGCGGACACTGATCAATTTTGA 961
DB 661 CGATCTCTGCGGACCGACACATCATTTGCTTCCGCGGACACTGATCAATTTTGA 720

QY 962 TACCACTACGAGCAATTCCTTGGCTAAAGAAAATGATACGCTACGCTACGCA 1021
DB 721 TACCACTACGAGCAATTCCTTGGCTAAAGAAAATGATACGCTACGCTACGCA 780

QY 1022 TTTGTTCAATGCGATGCTTCGCTGATCAATAGGCTCCCGGACGCGTGGGCTTTGCT 1081
DB 781 TTTGTTCAATGCGATGCTTCGCTGATCAATAGGCTCCCGGACGCGTGGGCTTTGCT 840

QY 1082 TGCTGCGGCACTGCGCGGAGACGATATGTTGATGCGGACGCGTGGATTTGCT 1141
DB 841 TGCTGCGGCACTGCGCGGAGACGATATGTTGATGCGGACGCGTGGATTTGCT 900

QY 1142 CGATGGAACGATGATGCTGTTTCAACAAAGCTTTTATCATCAAGACCGCATGA 1201
DB 901 CGATGGAACGATGATGCTGTTTCAACAAAGCTTTTATCATCAAGACCGCATGA 960

QY 1202 AGCGCGGGAATGCGAGACGAGTATCATTTTGGGCGGTTTGAAGTCAACGCGA 1261
DB 961 AGCGCGGGAATGCGAGACGAGTATCATTTTGGGCGGTTTGAAGTCAACGCGA 1020

QY 1262 TGAAGTCCGCTCTGCGGATGCGGCGCCATCCCGGAGGACACGACACTAGCGAG 1321
DB 1021 TGAAGTCCGCTCTGCGGATGCGGCGCCATCCCGGAGGATACGACACTAGCGAG 1080

QY 1322 TGAAGTCCGCTCTGCGGATGCGGCGCCATCCCGGAGGATACGACACTAGCGAG 1381
DB 1081 TGAAGTCCGCTCTGCGGATGCGGCGCCATCCCGGAGGATACGACACTAGCGAG 1140

QY 1382 AACCGTCCGCTCTGCGGATGCGGCGCCATCCCGGAGGATACGACACTAGCGAG 1441
DB 1141 AACCGTCCGCTCTGCGGATGCGGCGCCATCCCGGAGGATACGACACTAGCGAG 1200

QY 1442 AAATTTTGTGCTTTGACTCAAAACGCGGAGTGAAGATCAATTAAGTCAATCACT 1501
DB 1201 AAATTTTGTGCTTTGACTCAAAACGCGGAGTGAAGATCAATTAAGTCAATCACT 1260

QY 1502 ACTTTAGTACGATTAATCTATCTG 1528
DB 1261 ACTTTAGTACGATTAATCTATCTG 1287

RESULT 7
US-10-781-014-415
Sequence 415, Application US/10781014
Publication No. US20040180408A1
GENERAL INFORMATION:
APPLICANT: Pompeius, Markus
APPLICANT: Kroger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Habermeyer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
FILE REFERENCE: BGI-126CPN
CURRENT FILING DATE: 2004-02-17
PRIOR FILING DATE: US 09/602,740
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141,031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/143,208
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 60/151,572
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19931412.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931413.6
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931419.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931424.1
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931428.4
PRIOR FILING DATE: 1999-07-08
Remaining Prior Application data removed - See file wrapper or PALM.
SEQ ID NO 415
LENGTH: 1287
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: CDS
LOCATION: (101)..(1264)
OTHER INFORMATION: RXN00043
US-10-781-014-415

Query Match 20.7%; Score 1237.4; DB 18; Length 1287;
Best Local Similarity 97.6%; Pred. No. 0;
Matches 1256; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

242 AACAGAGGCTCAAGTCCGAGATTAATTAACCTAGATCCGTAGACATATGACATCATC 301
1 AACAGAGGCTCAAGTCCGAGATTAATTAACCTAGATCCGTAGACATATGACATCATC 60
302 GTCTTATGCTTCTGAGAGAAACCAATTAACCTAGAGATGCGAGAAAGTGTGATTA 361
61 GTCTTATGCTTCTGAGAGAAACCAATTAACCTAGAGATGCGAGAAAGTGTGATTA 120
362 TCAGAGAAATGAGAGTCAAGGATTAATTAATTAATGAGAGAAATTTGTTTCCCTCC 421
121 TCAGAGAAATGAGAGTCAAGGATTAATTAATTAATGAGAGAAATTTGTTTCCCTCC 180
422 GGTGATGAGTGTCTTCTCAACTCGAAAAAGGATCATACAGAACTCTGAGAAAC 481
181 GGTGATGAGTGTCTTCTCAACTCGAAAAAGGATCATACAGAACTCTGAGAAAC 240
482 AGCACTTAAAAAGCAGAGATTCCACCAGAACTCCCAAGATTGTTCCGATTATTTGA 541
241 AGCACTTAAAAAGCAGAGATTCCACCAGAACTCCCAAGATTGTTCCGATTATTTGA 300
542 TCTTATATACAGGAGTGAAGAGTGGCGCTTCTTCAAGGAGAGAGAGAGAGAGAG 601
301 TCTTATATACAGGAGTGAAGAGTGGCGCTTCTTCAAGGAGAGAGAGAGAGAGAG 360
602 GAACACCGGCGAGATTAACCGGAAACATGACAGACCGATGATGTCAGAGATGTTTC 661
361 GAATCGCGGCGAGTATACCGGAAACATGACAGACCGATGATGTCAGAGATGTTTC 420
662 GCGCGCGGCGAGTACGATGCGAGCGGAGTGAAGAACTTATTCCTTGTGAAGAGT 721
421 GCGCGCGGCGAGTACGATGCGAGCGGAGTGAAGAACTTATTCCTTGTGAAGAGT 480
722 CTGCTGTCGCGGATTAACCTGAGGAGCTTTCATCAAGGAGTGGTGGTGTCA 781
481 CTGCTGTCGCGGATTAACCTGAGGAGCTTTCATCAAGGAGTGGTGGTGTCA 540
782 AAACCCGAGATTTTATTTTCCGCGAACCCCAAGATCTTTCGCGGATCATGCGAG 841
541 AAACCCGAGATTTTATTTTCCGCGAACCCCAAGATCTTTCGCGGATCATGCGAG 600
842 AAAGGTTGATCAATGATCAAGTACGAGTGGCGGAACTGACATCTTTCGAGCTTCT 901
601 AAAGGTTGATCAATGATCAAGTACGAGTGGCGGAACTGACATCTTTCGAGCTTCT 660
902 CGATCTGCGGAGGAGCAGCATCATGCTTCTTCCGCGACACATGAGATGATTTTGA 961
661 CGATCTGCGGAGGAGCAGCATCATGCTTCTTCCGCGACACATGAGATGATTTTGA 720
962 TACCACTACAGCGCAATGCTTGGCTTAAGAGAAAAATGTACGGTCAACGGCTACCA 1021
721 TACCACTACAGCGCAATGCTTGGCTTAAGAGAAAAATGTACGGTCAACGGCTACCA 780
1022 TTTGTTCAATGAGATGCTTCCGCTGATCAATAGGCTCCCGGACGTTGGGCTTTTCT 1081
781 TTTGTTCAATGAGATGCTTCCGCTGATCAATAGGCTCCCGGACGTTGGGCTTTTCT 840
1082 TGCTCGGAGCAGTGGCGGAGGAGCATATGTTGATGCTGCGAGCGGAGTTCATTTGAC 1141
841 TGCTCGGAGCAGTGGCGGAGGAGCATATGTTGATGCTGCGAGCGGAGTTCATTTGAC 900
1142 CGATGAGAACGATGATCTAGCTGTTCCAAACAGCTTTTTCATACGGAGCCATGGA 1201
901 CGATGAGAACGATGATCTAGCTGTTCCAAACAGCTTTTTCATACGGAGCCATGGA 960
1202 AGCGCGCGGAGATGCGAGAGTGAATATTTTGGCGTTTGAACGTCACCGTACCGA 1261
961 AGCGCGCGGAGATGCGAGAGTGAATATTTTGGCGTTTGAACGTCACCGTACCGA 1020

1262 TGAGTCCGCTCTGCGGATGCGGCGGAGCCATGCGCGGAGGAGCCAGACATGAGAG 1321
1021 TGCGTCCGCTCTGCGGATGCGGCGGAGCCATGCGCGGAGGAGCCAGACATGAGAG 1080
1322 TCAGTTCGTCAGCAGTGGCGGAGGATGATGAGCTTATGAGCGGACCTTCCACCTC 1381
1081 TCAGTTCGTCAGCAGTGGCGGAGGATGATGAGCTTATGAGCGGACCTTCCACCTC 1140
1382 AACGTCGCGGATTAATTTCTGCTTGGCGATGAGAAATGCTTAATTCACCTTGC 1441
1441 AACGTCGCGGATTAATTTCTGCTTGGCGATGAGAAATGCTTAATTCACCTTGC 1200
1442 AAATTTGCTTCTGCTGATGAGAGGCGGAGTGAAGGATGATTAAGTCAATCAAGT 1501
1201 AAATTTGCTTCTGCTGATGAGAGGCGGAGTGAAGGATGATTAAGTCAATCAAGT 1260
1502 ACTTAAGTACAGATTAATTCCTG 1528
1261 AATTTAATACAGCAAACTTCTG 1287

RESULT 8
US-09-738-626-2907
; Sequence 2907, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIKOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 2907
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-2907

Query Match 18.6%; Score 1110.4; DB 9; Length 1152;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 1126; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

354 GTGCAATTATCAAGAAATGACAGTCAAGCACTTAATAAATTTGAGGAGAAATTTGCC 413
1 GTGCAATTATCAAGAAATGACAGTCAAGCACTTAATAAATTTGAGGAGAAATTTGCC 60
414 CCCCTCGGCGGATGATGAGTGTCTTCTCCAACTGAGAAACGAGCATCAAGGAACTCTCT 473
61 CCCCTCGGCGGATGATGAGTGTCTTCTCCAACTGAGAAACGAGCATCAAGGAACTCTCT 120
474 GAGAACACGACCTTAATAAAGCAGAGATTCCACCCGAACTCCCAAGATTGTTCCGGT 533
121 GAGAACACGACCTTAATAAAGCAGAGATTCCACCCGAACTCCCAAGATTGTTCCGGT 180
534 TTTATGATCTTCAATATCAAGTGAAGAGTGGCGCTTTCTTACGGAGAACGAGAGAC 593
181 TTTATGATCTTCAATATCAAGTGAAGAGTGGCGCTTTCTTACGGAGAACGAGAGAC 240

[illegible]

```

APPLICANT: Schroder, Hartwig
APPLICANT: Zeidler, Oskar
APPLICANT: Haberhauser, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
TITLE OF INVENTION: INVOLVED IN CARBON METABOLISM AND ENERGY
FILE REFERENCE: BGI-126CPCN
CURRENT APPLICATION NUMBER: US/10/781,014
CURRENT FILING DATE: 2004-02-17
PRIOR APPLICATION NUMBER: US 09/602,740
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141,031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/143,208
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 60/151,572
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: DE 19931412.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931413.6
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931419.5
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931420.9
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931424.1
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931428.4
PRIOR FILING DATE: 1999-07-08
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 784
SEQ ID NO 423
LENGTH: 882
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: CDS
LOCATION: (101)..(859)
OTHER INFORMATION: KXA00042
US-10-781-014-423

```

Query Match	12.8%	Score 762;	DB 18;	length 882;
Best Local Similarity	93.0%	Pred. No. 2e-223;		
Matches 820;	Conservative 0;	Mismatches 60;	Indels 2;	Gaps 2;
Qy 1449	GTGGTCTTTGACTCAACGGCGCAGGTGCMAAAGGTCCATTGAGTCAATCAAGTACTTTAA	1508		
Dp 1	GTGGTCTTTGACTCAACGGCGCAGGTTCMAACAGGTCCATTGAGCATCAAGTAAATTTTAA	60		
Qy 1509	GTACGAGTAAACCTATCTCTGAT-TTTTAAAGAGTCCCAACCATGGAATCACTATCTGCA	1567		
Dp 61	ATACGACMAAACCTTCTCTGATTAATAAAGAGTCGACCATGGAACATCAATCTGCA	120		
Qy 1568	AGACGACCAAGAAATGCGCAAGAGTGTGCAGTCTATCGCACCCCTTGCACAAAGG	1627		
Dp 121	AGACGACCAAGAAATCGGCAAAAGACGGCGAGCCCTGATCGCACCTTGCCACTPAAGG	180		
Qy 1628	TGAACTCTGGGGCTTGCAACAGGATCCTCAACACTAGTACTTACCAAGAGTCAATTCG	1687		
Dp 181	CGGAACCTTGGGGCTTGCAACTGGATCGTCACTTTGAGCACTTACCAAGAGTCAATTCG	240		
Qy 1688	CATGTATGAAGCTGGGGAGTGTCAATTCAGAACTGCAAGGCAATTCCTGTGATGATGA	1747		
Dp 241	CATTTATGAAGCTGGGGAAGTGTCAATTCAGAACTGCAAGGCAATTCCTGTGATGATGA	300		
Qy 1748	CGTGGGACTTAAACCCGTCAGATGAAAAACAGTACTTTTAAAAACATTCGCAAAAGAGTTCAC	1807		
Dp 301	CGTGGGACTTAAACGGCGACAGATGAAAAACAGTACTTTTAAAAACATTCGTAAAGAGTTCAC	360		
Qy 1808	TGACCAATCGAACAATCGTTGATGGAAGAGTCTAAGCCCAAGATGTGCAAAACCGTAGTCC	1867		
Dp 361	TGACCAATCGAACAATCGTTGATGGAAGAGTCTAAGCCCAAGATGTGCAAAACCGTAGTCC	420		

QY 1868 ATACGAGAGAGCTGAGATGATGAGCAAGATCGCTGAGATCCGTTGAAGTTCAAT 1927
DB 421 ATACGAGAGAGCTGAGATGATGAGCAAGATCGCTGAGATCCGTTGAAGTTCAAT 480
QY 1928 CTTGGGATCGGCGGAAAGG-CACATCGCTTTCATTTGAACCATCATCTTCTCTGAG 1986
DB 481 CTTGGGATCGGCGGAAAGGCGGCAATCGCTTTCATTTGAACCATCATCTTCTCTGAG 540
QY 1987 ACTGACAAAGTCCAGGCGCTGACCCCTTAAACTGTGAGGACAAAGCTGATTTCTCA 2046
DB 541 ACTGACAAAGTCCAGGCGCTGACCCCTTAAACTGTGAGGACAAAGCTGATTTCTCA 600
QY 2047 CACCATGAAAGAGTCCCAACCCAGCGCTGACCCAGGTTGGGCACTTTTCCGCGC 2106
DB 601 CACCATGAAAGAGTCCCAACCCAGCGCTGACCCAGGTTGGGCACTTTTCCGCGC 660
QY 2107 GCAAAACATCGTTGGTGGCACTGTGTAAGAAAGCCGACGCCATCGCGGAACTGT 2166
DB 661 GCAAAACATCGTTGGTGGCACTGTGTAAGAAAGCCGACGCCATCGCGGAACTGT 720
QY 2167 GGAAGGCCAGTACTGCTTTTGGCCAGGTTCCATCCCTGTAGTCAACAATGCGCAC 2226
DB 721 GGAAGGCCAGTACTGCTTTTGGCCAGGTTCCATCCCTGTAGTCAACAATGCGCAC 780
QY 2227 ATCATCGTTGATGAAGAGCAGATGATCAAGCTGAAAGCGCTGATCACTACGCTCAT 2286
DB 781 ATCATCGTTGATGAAGAGCAGATGATCAAGCTGAAAGCGCTGATCACTACGCTCAT 840
QY 2287 GAGCAATTAAGCTGCGCTGAGAAACAAAGAAAGTACTG 2328
DB 841 GAGCAATTAAGCTGCGCTGAGAAACAAAGAAAGTACTG 882

RESULT 10
US-09-738-2906
; Sequence 2906, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 2906
; LENGTH: 759
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-2906

Query Match 11.3%; Score 671.8; DB 9; Length 759;
Best Local Similarity 93.7%; Pred. No. 1.3e-195;
Matches 711; Conservative 0; Mismatches 47; Indels 1; Gaps 1;
QY 1548 ATGGAATGACTATCTGCAAGAGAGAGAGAGTGGCAAGAGCTTCACTGCTAATC 1607
DB 1 ATGGAATGACTATCTGCAAGAGAGAGAGAGTGGCAAGAGCTTCACTGCTAATC 60

QY 1608 GCACCTTCGCCAAGAGGTGGAACCTTGGGGCTTGGAAAGATCCCTACACTGAGT 1667
DB 61 GCACCTTCGCCAAGAGGTGGAACCTTGGGGCTTGGAAAGATCCCTACACTGAGT 120
QY 1668 ACTTACCAAGAGCTATTCGATGTATGAGTGGGGAAGTGTCAATTAAGAGTCAAG 1727
DB 121 ACTTACCAAGAGCTATTCGATGTATGAGTGGGGAAGTGTCAATTAAGAGTCAAG 180
QY 1728 GATTCCTTGTGATGAATACGTGGGACTAACCCGTGACGATGAAGAAAGAGCTTAA 1787
DB 181 GATTCCTTGTGATGAATACGTGGGACTAACCCGTGACGATGAAGAAAGAGCTTAA 240
QY 1788 ACCATTGCAAGAGTCACTGACCAATCGACATCGTTGATGAAGAGTTCAGAGCCA 1847
DB 241 ACCATTGCAAGAGTCACTGACCAATCGACATCGTTGATGAAGAGTTCAGAGCCA 300
QY 1848 GATGTGCAAAACCTGATTCATCGAAGAGTGGAGATGAGGCAAAAGTCTGCA 1907
DB 301 GATGTGCAAAACCTGATTCATCGAAGAGTGGAGATGAGGCAAAAGTCTGCA 360
QY 1908 GAATCCGTTGAAGTCAAAATCTTGGCATCGCGGAAACGG-CACATGCTTTCATGAA 1966
DB 361 GAATCCGTTGAAGTCAAAATCTTGGCATCGCGGAAACGGCACATGCTTTCATGAA 420
QY 1967 CCATCATCTTCTGTGCAAGAGTCCAGAAAGTCCAGGCGCTGACCCCTAAACTGGAG 2026
DB 421 CCATCATCTTCTGTGCAAGAGTCCAGAAAGTCCAGGCGCTGACCCCTAAACTGGAG 480
QY 2027 GACAAAGCTGCAATTCCTCAACACCATGCAAGAGTCCCAACCCAGCGCTCACCCAGGT 2086
DB 481 GACAAAGCTGCAATTCCTCAACACCATGCAAGAGTCCCAACCCAGCGCTCACCCAGGT 540
QY 2087 TTGGGCACTTTTCCCCGGCGGAAACATCGTTGTGGCAACTGTGGAAGAAAGCC 2146
DB 541 TTGGGCACTTTTCCCCGGCGGAAACATCGTTGTGGCAACTGTGGAAGAAAGCC 600
QY 2147 GAGCCATCCGCGGAAAGTGGAAAGGCCAGTACTGCTTCTTGGCCAGGTTCCATCTCG 2206
DB 601 GAGCCATCCGCGGAAAGTGGAAAGGCCAGTACTGCTTCTTGGCCAGGTTCCATCTCG 660
QY 2207 TAGATGCAAAACATGCAACATCATGTTGATGAAGAGAGATTCAGAGTGAAGAAAC 2266
DB 661 TAGATGCAAAACATGCAACATCATGTTGATGAAGAGAGATTCAGAGTGAAGAAAC 720
QY 2267 GCTGATCACTACCGTCTCATGAGCAATTAAGCTGCGC 2305
DB 721 GCTGATCACTACCGTCTCATGAGCAATTAAGCTGCGC 759

RESULT 11
US-10-494-541-9
; Sequence 9, Application US/10494541
; Publication No. US20050009152A1
; GENERAL INFORMATION:
; APPLICANT: Zelder, Oskar
; APPLICANT: Pompeius, Markus
; APPLICANT: Schroder, Hartwig
; APPLICANT: Kroger, Burkhard
; APPLICANT: Klopptroge, Corinna
; APPLICANT: Haberer, Gergor
; TITLE OF INVENTION: Gene coding for proteins for genetic stability,
; FILE REFERENCE: BGI-166US
; CURRENT APPLICATION NUMBER: US/10/494,541
; CURRENT FILING DATE: 2004-05-03
; PRIOR APPLICATION NUMBER: PCT/EP02/12138
; PRIOR FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: DE 10154180
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 104
; SEQ ID NO 9
; LENGTH: 1543

TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: CDS
LOCATION: (101)..(1513)
OTHER INFORMATION: RXA00314
US-10-494-541-9

Query Match
Best Local Similarity 100.0%; Pred. No. 8.4e-79;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5674 ACCGTAAACACTAGGTTTGGGCGAAATGAAAGCGGCAACCTGCTCAACGTCG 5733
DB 1 ACCGTAAACACTAGGTTTGGGCGAAATGAAAGCGGCAACCTGCTCAACGTCG 60
QY 5734 AAAGAAAGAGCGGTCGCAACAACCATTAAGTTGAAACCTTAGTTGCAACAGT 5793
DB 61 AAAGAAAGAGCGGTCGCAACAACCATTAAGTTGAAACCTTAGTTGCAACAGT 120
QY 5794 TAGACTAGGAGAGCTGACTTACGCATCTTTGACACCGGTAACCGCTTGAGATT 5853
DB 121 TAGACTAGGAGAGCTGACTTACGCATCTTTGACACCGGTAACCGCTTGAGATT 180
QY 5854 TAAACCTGTTCAACGAGTATGCTCGGTGTAAGTGTGAGCAACCGCGAATCTTC 5913
DB 181 TAAACCTGTTCAACGAGTATGCTCGGTGTAAGTGTGAGCAACCGCGAATCTTC 240
QY 5914 ACCCCACATTTGACATGTTGTTGAGCAGTACGCGTTGATTTTGGCGCTG 5967
DB 241 ACCCCACATTTGACATGTTGTTGAGCAGTACGCGTTGATTTTGGCGCTG 294

RESULT 12
US-09-746-660A-31/C
Sequence 31, Application US/09746660A
Publication No. US20030049804A1
GENERAL INFORMATION:
APPLICANT: Pompejus, Markus
APPLICANT: Krieger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauer, Gregor
APPLICANT: Kim, Jun-Won
APPLICANT: Lee, Heung-Schick
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
FILE REFERENCE: BGI-121CP2
CURRENT APPLICATION NUMBER: US/09/746,660A
CURRENT FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 09/606740
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 09/603124
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/142101
PRIOR FILING DATE: 1999-07-02
PRIOR APPLICATION NUMBER: 60/148613
PRIOR FILING DATE: 1999-08-12
PRIOR APPLICATION NUMBER: 60/187970
PRIOR FILING DATE: 2000-03-09
PRIOR APPLICATION NUMBER: DE 19931420.9
NUMBER OF SEQ ID NOS: 125
SOFTWARE: PatentIn Vers. 2.0
SEQ ID NO 31
LENGTH: 1059
TYPE: DNA
ORGANISM: Corynebacterium glutamicum
FEATURE:
NAME/KEY: CDS

LOCATION: (101)..(1036)
OTHER INFORMATION: RXA00044
US-09-746-660A-31

Query Match
Best Local Similarity 96.6%; Score 215.2; DB 10; Length 1059;
Matches 220; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 AGTCGTCAGACGCCACCATTTGATGTGTGTGTCACCGAGCTTGCGAGGCTTTCACTT 60
DB 228 AGTCGTCAGACGCCACCATTTGATGTGTGTGTCACCGAGCTTGCGAGGCTTTCACTT 169
QY 61 ACCCTCCCGTCGCGCGTGGAGTGGGTCATTACCGGTGGATCAACCGCGTGAAGTTGCG 120
DB 168 ACACCTCCCGTCGCGCGTGGAGTGGGTCATTACCGGTGGATCAACCGCGTGAAGTTGCG 109
QY 121 GAACCATGTTGTTCTTGTGGTTGAGGAAAGAGTGGCGGTGAGAAATTTTCAAGTG 180
DB 108 GAACCATGTTGTTCTTGTGGTTGAGGAAAGAGTGGCGGTGAGAAATTTTCAAGTG 49
QY 181 TCTGCAATTTTAAATTATGATCATCATGCTTGAAAGCTGAGTAAT 228
DB 48 TCTGCAATTTTAAATTATGATCATCATGCTTGAAAGCTGAGTAAT 1

RESULT 13
US-10-282-122A-17961
Sequence 17961, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Lianguu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See file Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 17961
LENGTH: 777
TYPE: DNA
ORGANISM: Corynebacterium diptheriae

US-10-282-122A-17961

Query Match	3.5%	Score 208;	DB 17;	Length 777;
Best Local Similarity	57.8%;	Pred. No. 1.9e-52;		
Matches 428; Conservative	0;	Mismatches 305;	Indels 7;	Gaps 31

QY	1548	ATGGAATACACTATCTGGCAAAAGACAGACAAGAAAGTCGGCAAAAGCAGTTGACGTCTAATC	1607
Db	1	ATGGAGATCTGTAATACCCGCCCAACAGACGACGCCCTCGTATGCCGCCGACATCTTA	60
QY	1608	GCACCTTCGCGCAACAGGGGTGAACTTGGGGCTTGGCAACAGATCTCTCAACCTAGAT	1667
Db	61	GAAAGATACGACACGTGAAGAAAGACCCTGGGTCTTGGTACCGGATCTAATCTCGCTAAGC	120
QY	1668	ACCTTACCAAGCTCTATTCGCATGTATGAGCTGGGGAAAGTCTCATTTCAAGATCTGAAG	1722
Db	121	ACCTAACCAAGCTCATTCGTGCGCCACATGAGGAAGGCTTGTCTTTTGGCGGAATGCCAA	180
QY	1728	GCATCTTGTGGATGAATACGTGGGACCTAACCCGTGACGATGAAACAGCTACTTTAA	1787
Db	181	GCATTTACTCTCGACGAATACGTCCGATTTGCTCTCGCAACAGAACAAAGCTACTACAC	240
QY	1788	ACCATTCGCAAAAGTTCCTGACCACTGCATCATCGTTGATGAAAGGTCTACAGCCCA	1847
Db	241	ACCATTCGACGGGAATTTACTTCGATATCGATATCCAGACGAAAGTGTTCAAACCT	300
QY	1848	GATGTGCAAACTCTGATCCATACAGACAGCTGCAGATGATGAGGCAAAAGTGTCTCA	1907
Db	301	GACGGCACCGGAGAACCCCGAGACACAGACACAGAAATACGATCGCTGATCGTAGAA	360
QY	1908	GAATCC---GTGGAAGTCAAAATCCTTGGACATCCGGCGGAAACCG-CAATCGCTTTCAT	1963
Db	361	AAAGCGGTGTGACATTCAGATTTCTTGATTCGAAACGATGGTCAATGCTTTTAAAC	420
QY	1964	GAACCATCATCTTCTCTGTGACGACTGCAAAAGTCCAGGGCTGTCAACCTTAAACTGTG	2023
Db	421	GAGCCAACTGTGTATGAGCAAGCGGTACGGGTATCAAGAGCGTGTACACCAGACAGGTA	480
QY	2024	GAGGCAACGTCGATTTCTTCAACACCATGCA---AGAGTCCCAACCCAGCGCGTCAAC	2080
Db	481	CGCATATATTTCTGATTTCTTTAAATGGCAACGAATCACAAGTTTCAACACAGTATGAC	540
QY	2081	CAGGTTTGGGCACTTTGTCCCGGCGGCAAAACATCGTGTGTGGTGCAACTGTGGAAGA	2140
Db	541	CAGGAATCGGCAACCATTCGCGAGGCCGACACCTGTGATGCTCTGTGTTGGGAGAAAT	600
QY	2141	AAAGCCGACGCGCATCCGCGGAACGTGTGAAGGCCCAGTACGTCTTTCGGCCAGTTC	2200
Db	601	AAACGATGTGAGTCAAAAGCATTGTGGAAGGGCAATGTGCAAGCGGTGTCCAGCATCT	660
QY	2201	ATCTGTATGATGCAACAACATGCCACCATCATCGTTGATGAAGACAGATATCCAAAGCTG	2260
Db	661	GTTTTGAAGCTGCATGAGCATGCCACTGTGATTTGTGATGAGGAGCGGCGAGCAAACTT	720
QY	2261	GAAGAACGTATGACTACCG 2280	
Db	721	GAGCACTACGTACTACCG 740	

RESULT 14
US-10-282-122A-17959

? APPLICANT: Wang, Liangsu
 ? APPLICANT: Zamudio, Carlos
 ? APPLICANT: Malone, Cheryl
 ? APPLICANT: Haselbeck, Robert
 ? APPLICANT: Ohlson, Kari
 ? APPLICANT: Zykkind, Judith
 ? APPLICANT: Wall, Daniel
 ? APPLICANT: Trawick, John
 ? APPLICANT: Carr, Grant

```

/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Foreyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 17959
/ LENGTH: 1212
/ TYPR: DNA
/ ORGANISM: Corynebacterium diptheriae
US-10-282-122A-17959

```

Query Match	3.2%	Score 192.6	DB 17	Length 1212
Best Local Similarity	53.5%	Pred. No. 1.4e-47		
Matches 433	Conservative 0	Mismatches 364	Indels 12	Gaps 1
Qy	511	AACGCCACGATGTTCCCGGTTTATATGATCTTCATATACAGGTGAAACGGTGGC	570	
Db	194	ACCACCCCTCGTCTGCAAGTCTTGGCCGACATCCAAATCAGGTGACAGAGGAGT	253	
Qy	571	CGTTTCTACGGGAACGACGACGAGGAGAAACCCGCCAGTATCACCGCGACATG	630	
Db	254	CCTTCCCACTCCGACTACGACGAGTGCAGTCCGCGCACAGCACACCGCGCACATG	313	
Qy	631	GCACGACCGGATGTTGGCAAGCATGATTTCCGGGCGCGGCGTGAAGCACTGGCAACGGAG	690	
Db	314	GCTGACACAGCTCTGAGCTGCGACAGATTCATATGCCAGACACACCTTACTGCGCAAC	373	
Qy	691	TGGAACAACCTTATTCCTTGTGTGTAAGAGCTCTGCTGTGGGCAATTCACCTCGAGGACC	750	
Db	374	TATCCCTCTAGCGGATCTTGGCCGACCCCGTGAAATTCACGAAATCAGCCGGAAGCC	433	
Qy	751	CTTTCATCAACGATCGATCGCTTGTGTGCTCAAAACCCGGATTTCAATTTTCCGGCAACC	810	
Db	434	CCTTGTGTAACCGGTGCGCTGCGGCGCACAAAGCCAGAAAGCATATCTTAGCGCAC	493	
Qy	811	CAACAGATCTTGTCCCGGGTATTCATATCGGGAAGGTTGGATTAATTCATCACTAG	870	
Db	494	CAGAACTGTTCAAGAAATATGACAGCGCACCGGCTGGCTCAAAATCCATGACCTTGC	553	
Qy	871	CGCGGAAACTGACAATCTTTCTGAGCTTCCGATCTCTGGCGAGCGCACCACTCATTTG	930	
Db	554	CCCCGAAACCGCACGCGCAAGGAAATCATCGACTCTTGCGCGGAAACCAACTCATTCG	613	
Qy	931	CTTCTTTCGGGCACTATGATGAGATTTTGATACCACTACGACGCAATTTGCTTGGCTA	990	
Db	614	TCTCCCTCGGACACACGAGCGAGACTTCTCAGTCAACGAAACCAAGCCCTGTCTTACGAG	673	
Qy	991	AAGAGAAAAATGTGACGTCACGGCTACGCAATTTGTTCAATGCGATGCTTCGCTGATC	1056	

Db 674 TAGGAGCGCGTGCAGTGTGACCGCAACCCACCTATTATGCAATGCCAGCAATPACAC 733
Qy 1051 ATAGGGCTCCCGGACGCGTGGCGCTTGTCTGTGCGGACGCGCGGGAGCATATG 1110
Db 734 ACCGTGACCCCGGTGACGCGGACGCGCTTATCGACCCGACGTCGAGCAACGCCACG 793
Qy 1111 TTGAGTGTATGCGCGGCGGTGCAATTTGGCCGATGGAACG-----TCGATC 1158
Db 794 TAGAACTAGTCGCGCATGCGATCCACTGTGATGACACAGCGCGCATGTGATGACT 853
Qy 1159 TAGCTGTTCACAAACGCGCTTTTCATGACGAGCGCCATGGAAGCGCGGAAATGCCAG 1218
Db 854 CGGTAGTGGCGATGCAAGTGAAGCTTGTCTCGACCGCATGGGCGAGAGAAAGAG 913
Qy 1219 ACGGTGAGTACATTTTGGCGCTTTTGAAGTCACCGTCAACGATGAGTGGCGCTTGC 1278
Db 914 ACGGGGACTACCTTCTCGGAGCGCTGCGCGTCAACGTAAGACTCCGTCGACGACTGA 973
Qy 1279 GGGATGGCGCGCATGCGCGGGGGGACG 1307
Db 974 CACCAACGAGCGGAGGAGGCGCTATC 1002

RESULT 15
US-09-974-300-829
; Sequence 829, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 829
; LENGTH: 975
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-829

Query Match 3.2%; Score 192; DB 9; Length 975;
Best Local Similarity 51.4%; Pred. No. 1.9e-47;
Matches 472; Conservative 0; Mismatches 440; Indels 6; Gaps 1;

Qy 4240 TGTTCGTGAGATGATCAACCTGATGCGCGCTTCTTGCAGTGTGTG 4299
Db 27 TCTCGCTAATGATGATTAATCTGATGCGGGAATGCTTTCGATTTCTCCGCGCTAAT 86
Qy 4300 TGTTCACCGCAACCAACGCTTGTGCGCAATGAGTTCGCGCGCGCGCATTTGGTAT 4359
Db 87 CGGCTGTGTGCGCGCTCAACGCGTGTGGCGAAATCGGCTGCTCGGCTCGGCGCT 146
Qy 4360 GCGGATGTGTGTTCCCAACCTGCTTAAGCGGCTAAGCAAGTGGCGCGCACATGACGCGGG 4419
Db 147 CATGCTTGTGACCCCTGATTTGTTAAACCGCTGGGATACGGGGCTGCGAGCAAAAGCGG 206
Qy 4420 CGAAATGCCAATGTGTCCCTGTTTGGTGTGATGTTGCTCAAGCTGTACAGGCGCAC 4479
Db 207 AAGAGATCTGTCTGGAATTTATGCGTTTGAAGTGAAGAGTGTGCTACAGGGGCA 266
Qy 4480 CGTGTCTTGTGTGTGTGTGTCTTGTGATTTCTGCAACGATGAGAGTTCCTGCAAA 4539
Db 267 GGTGTCTTCCATTTTGT 326
Qy 4540 GCGACTCATGGGACATGACAGACTTCTGATCAACCCAGTGTGATCTGTGCTACCGG 4599

Db 327 GGGAGCGCTGAGAGACATTCAGCTCTCTGTGTGGACCGATTAACGCTCTGTGACAGG 386
Qy 4600 TTTCCTTAAGTTGATTTGATTTGTCACGAATGCGTGGGTGATCTGCGACA 4659
Db 387 TTTCCTTCAATTATTTGCTTGAACCGATTAATTTCCATTCGGAAATGTTTAAACATC 446
Qy 4660 CGGTGTGAGGAGACTATGATTTGCGGTGTCATGCGCGGTGTGCTTTGCGTGTGT 4719
Db 447 AGGCTTATTTGTGCTTTGAACAATTTGCGGTACTGCGGAGCTGTTATACGAGGGCT 506
Qy 4720 CTACTCAACAAATGTTATCACTGTGTGACCAAGCTCTTCCGCAATTTGAGCTGAGCT 4779
Db 507 TTACCGCGCTCTGTCATTAACGGAATGATATATTTCTGTGCGCTTGAACCTTCACT 566
Qy 4780 GTTC-----AACAGGATGATCTTCACTTTCGCAACCGCATCATGCGCAATATGCG 4833
Db 567 GATCGGCTCAAAAGCTCGCGGAAATTTTATGCGGATGCTGCGGCTGTCCAAATATGCG 626
Qy 4834 GGAAGGTGACAGATTTTGGCAGTGTCTTCTAGCGAAGAGTGAAGAAAGTCAAGGGCT 4893
Db 627 ACMAAGTTCAAGCGCGCTTGCATGATGTTATGTTCAAGATGGAAGAGCAAGAAAGGCT 686
Qy 4894 TGCAGTGTCTTCAAGTGTCTCGCTGTTCTTGTATTAACAAGCTGCGATCTTGTGTGT 4953
Db 687 GTTCCTGACATCCGAATTTGACTTATCTCGGAATTAACGAACCGCATTTGTGAGT 746
Qy 4954 GAACCTTGCCTGCGCGCTTCTTCAATTTGATGATGATGATGATGATGATGATGATGAT 5013
Db 747 GATCTCAAGATACAGATTTTCCGTTGATGATGATGATGATGATGATGATGATGATGAT 806
Qy 5014 TTTGATGACCTTTGATATCAAGGCAAGTTCGTTGGCGCTGCAAGTTCCTTGTGTGT 5073
Db 807 GTTCATTTCTCGCAAGGGGTTTGGCAAGCTGTGCGCGCTGCGCGCGTACCTGGAAT 866
Qy 5074 TGTTCCTAATGATGCTCCAGATATGATGATGATGATGATGATGATGATGATGATGAT 5133
Db 867 TTTCTCTATTATGACCAATTAATGAGAGCGTTTGCATCGGAATGCGATGCTGTGAT 926
Qy 5134 CATGCAATTCGCGCGAGC 5151
Db 927 CGCGCGTTTGGCGGAGC 944

Search completed: March 9, 2005, 20:15:11
Job time : 2997 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 19:13:49 ; Search time 16576 Seconds
(without alignments)
13706.903 Million cell updates/sec

Title: US-10-019-284B-1

Perfect score: 5969
Sequence: 1 agtcgcgcgcgcgcgcacatt.....tgatatttgcgcgcgtgaa 5969

Scoring table: IDENTITY_NTC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_g881:*
9: gb_g882:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	174.8	2.9	917	8	BZ685850 PUBEB66TD
2	132.4	2.2	770	7	CN823265
C 3	123.6	2.1	890	6	CD499281 CDA38-D12
4	123.6	2.1	1030	6	CD499280 CDA38-D12
5	113.4	1.9	745	5	BP705177 BP705177
6	112	1.9	732	7	CK807155 AGENCOURT
7	111.8	1.9	760	7	CK805715 AGENCOURT
8	105.4	1.8	899	5	EX714658
9	104.4	1.7	811	7	CF346432 AGENCOURT
10	104.4	1.7	896	7	CF592490 AGENCOURT
11	104.4	1.7	1177	3	CR696680 Tetradon
12	104.4	1.7	1178	3	CR690088 Tetradon
13	104.4	1.7	1206	3	CR701935 Tetradon
14	103.8	1.7	818	4	BM017859
15	103.2	1.7	662	4	AL656223
16	103.2	1.7	902	7	CF378888 AGENCOURT
17	102.8	1.7	1173	3	CR703987 Tetradon
18	102.8	1.7	1160	3	CR699192 Tetradon
19	102.8	1.7	1193	3	CR698198 Tetradon
20	102.8	1.7	1203	3	CR674049 Tetradon
21	102.8	1.7	1226	3	CR669603 Tetradon
22	102.8	1.7	1242	3	CR665914 Tetradon
23	102.6	1.7	818	5	EX458246
24	102.6	1.7	846	7	CO572502 AGENCOURT

25	102.6	1.7	870	9	AY416649 Homo sapi
26	102.6	1.7	832	5	BX374555 BX374555
27	102.6	1.7	1033	1	AL560879 AL560879
28	102.6	1.7	1049	5	BX406188 BX406188
29	102.6	1.7	1051	4	BM476096 BM476096
30	102.6	1.7	1063	5	BM910583 AGENCOURT
31	102.6	1.7	1114	3	BX421863 BX421863
32	102.6	1.7	2229	3	CR609510 full-1eng
33	102.4	1.7	1040	9	AL152142 Anopheles
34	102.2	1.7	669	2	BE790435 BE790435
35	102.2	1.7	847	5	BU171522 AGENCOURT
36	101.8	1.7	1003	1	AL546345 AL546345
37	101.4	1.7	949	5	BX324827 BX324827
38	101.2	1.7	892	3	CR713846 Tetradon
39	101.2	1.7	917	5	BX428009 BX428009
40	100.2	1.7	875	5	BU158666 AGENCOURT
41	100	1.7	746	4	BI758774 BI758774
42	100	1.7	812	4	BI825295 BI825295
43	100	1.7	905	5	BU179037 AGENCOURT
44	100	1.7	934	5	BO929688 AGENCOURT
45	100	1.7	984	5	BU167870 AGENCOURT

ALIGNMENTS

RESULT 1
BZ685850/c 917 bp DNA linear GSS 05-FEB-2003
LOCUS PUBEB66TD ZM 0.6 1.0 KB Zea mays genomic clone ZMBETA035K11,
DEFINITION genomic survey sequence.
ACCESSION BZ685850
VERSION BZ685850.1 GI:28245462
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 917)
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org.
Seq primer: TP
Class: sheared ends.
Location/Qualifiers
1..917
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMBETA035K11"
/clone_lib="ZM 0.6 1.0 KB"
/note="Vector: pCR4-TOPO; Site_1: EcoRI; 0.6-1.0 kb high
cor selected genomic DNA library"

ORIGIN

Query Match 2.9%; Score 174.8; DB 8; Length 917;
Best local similarity 52.7%; Pred. No. 3.4e-39;
Matches 483; Conservative 0; Mismatches 412; Indels 21; Gaps 4;

QY 3930 TGAAGGCACTTGAACCTGCGGCGATGTCATTCGCGGCGGCGATGTGG 3989
DB 914 TCAAGGCTTCCTTCACCGGCGGCGTGTCCAGGTGTAATCGCCCGGTAAGTGG 855

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
Db	695	GGCGGCCCATGCTGGCGGGCCATTTTCGGCTCTGTACGACCGCTGGTGCATACCGG	754				
Db	4744	TCGCAACCAAGTCCTTC	4759				
Db	755	CATGCACCAACATGTC	770				
RESULT 3	CD499281/c	890 bp mRNA linear EST 12-JUN-2003					
LOCUS	CD499281	CD499281					
DEFINITION	CD499281	CD499281					
ACCESSION	CD499281	CD499281					
VERSION	CD499281.1	GI:31426312					
KEYWORDS	EST.						
SOURCE	Gasterosteus aculeatus	(three spined stickleback)					
ORGANISM	Gasterosteus aculeatus						
REFERENCE	1	(bases 1 to 890)					
AUTHORS	Kingsley, D.M., Peichel, C., Balaband, S., Grimwood, J., Dickson, M., Schmutz, J. and Myers, R.M.						
TITLE	Expressed sequence tags from Gasterosteus aculeatus						
JOURNAL	Unpublished (2003)						
COMMENT	Contact: Kingsley, DM HHMI and Department of Developmental Biology Stanford University School of Medicine Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5529, USA Tel: 650 725 5954 Fax: 650 725 7739 Email: kingsley@cmgm.stanford.edu Plate: 38 High quality sequence start: 13 High quality sequence stop: 799. Location/Qualifiers 1. 890 /organism="Gasterosteus aculeatus" /mol_type="mRNA" /strain="Salinas river, CA" /db_xref="taxon:69293" /clone="CD498-D12" /sex="mixed male and female" /tissue_type="heads and internal organs combined" /dev_stage="adult" /clone_id="SHGC-CD4" /note="Vector: Lambda ZAP Express/PBK-CMV; Site 1: EcoRI (5' adaptor); Site 2: XhoI (3' linker primer); The mixed organ cDNA library was generated using the ZAP-CDNA method by Stratagene. First strand cDNA synthesis was primed with a 50 bp linker primer containing an oligo dt sequence preceded by a synthetic XhoI site. 5 prime adaptors were used containing an EcoRI cohesive end. The finished cDNAs were inserted in to the ZAP express vector unidirectionally in the sense orientation with respect to the lacZ promoter of PBK-CMV. An amplified library was prepared from approximately 3 million primary clones in the lambda ZAP Express vector. In vivo excision was then used to generate individual PBK-CMV phagemid clones for EST sequencing."						
FEATURES	Source						
ORIGIN	Query Match	2.1%	Score 123.6;	DB 6;	Length 890;		
	Best Local Similarity	52.3%	Pred. No. 4.7e-24;				
	Matches 346;	Conservative 0;	Mismatches 309;	Indels 7;	Gaps 3;		
Db	1606	TCGCAACCTTGCCCAAGGATGGAACCTTGGGGCTTGAACAGATCTTACCACTGA	1665				
Db	771	TCCGACCTGTCGCGACAGATTTTTCACCTGGGGCTTCCACAGAGAACCCCATAG	712				
Db	1666	GTACCTACCAAGAGCTCAATTCGATGTATGAAGCTGGGGAAGTGCATTCAGAAGCTGA	1725				

Accession	Version	KeyWords	Organism	Reference	Authors	Title	Journal	Comment
CD499280	1	GI:31426311	Gasterosteus aculeatus (three spined stickleback)	1961	Kingsley, D.M., Peichel, C., Balabhadra, S., Grimwood, J., Dickson, M., Schmutz, J., and Myers, R.M.	Expressed sequence tags from <i>Gasterosteus aculeatus</i> unpublished (2003)	HMNI and Department of Developmental Biology Stanford University School of Medicine Beckman Center B300, 279 Campus Drive, Stanford, CA 94305-5329, USA	<p>GenBank accession: CD499280</p> <p>Gene: <i>Gasterosteus aculeatus</i></p> <p>EST: 111 TG 110</p> <p>EST: 2259 TG 2260</p> <p>EST: 111 TG 110</p>

```

source
1..1030
/organism="Gasterosteus aculeatus"
/mol_type="mRNA"
/strain="Salinas river, CA"
/db_xref="taxon:69293"
/clone="CDA38-D12"
/sex="mixed male and female"
/tissue_type="heads and internal organs combined"
/dev_stage="adult"
/clone_lib="SHGC-CDA"
/notes="Vector: lambda ZAP Express/BK-CMV; site 1: EcoRI
(5' adaptor); site 2: XhoI (3' linker primer); The mixed
organ cDNA library was generated using the ZAP-CDNA method
by Stratagene. First strand cDNA synthesis was primed with
a 50 bp linker primer containing an oligo dT sequence
preceded by a synthetic XhoI site. 5 prime adaptors were
used containing an EcoRI cohesive end. The finished cDNAs
were inserted in to the ZAP express vector
unidirectionally in the sense orientation with respect to
the lacZ promoter of PBK-CMV. An amplified library was
prepared from approximately 3 million primary clones in
the lambda ZAP Express vector. In vivo excision was then
used to generate individual PBK-CMV phagemid clones for
EST sequencing."

```

ORIGIN

```

Query Match      2.1%; Score 123.6; DB 6; Length 1030;
Best Local Similarity 52.3%; Pred. No. 5e-24;
Matches 346; Conservative 0; Mismatches 309; Indels 7; Gaps 3;

```

```

1606 TCGACACCTTCGCCCAACAAGGTGGAACCTTGGGGCTTGCACAGAGATCCTCACCACGTA 1665
108 TCGACACCTTCGCCCAACAAGGTGGAACCTTGGGGCTTGCACAGAGATCCTCACCACGTA 167
1666 GTAACCTTCAAGAGCTCATTCGATGTAAGAGCTGGGGAGTGTCTTCAAGAACTGCA 1725
168 GTTGTATCAAGAACTAATCGAGTACTACAGATGGGCCAAATCTCATTTAGTATGTA 227
1726 AGGCAATCTTGTGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 1785
228 AAACCTTCAAGAGCTCATTCGATGTAAGAGCTGGGGAGTGTCTTCAAGAACTGCA 287
1786 AAACCTTCAAGAGCTCATTCGATGTAAGAGCTGGGGAGTGTCTTCAAGAACTGCA 1845
288 CTTGATGTAAGAGCTCATTCGATGTAAGAGCTGGGGAGTGTCTTCAAGAACTGCA 347
1846 CAGATGTAAGAGCTCATTCGATGTAAGAGCTGGGGAGTGTCTTCAAGAACTGCA 1902
348 TTGATGTAAGAGCTCATTCGATGTAAGAGCTGGGGAGTGTCTTCAAGAACTGCA 407
1903 CTGAGAAATCCGTTGAAGTCAATCTTGGGATGGGCGGAAACGG-CACATCGCTTCA 1961
408 CTGCGAGAGGATCGAGCTCTTGTGGAGTATTTGACACAGATGCGCCATTTGCTTCA 467
1962 TTGAACCTATCATCTTCTGTGAGAGCTGACAAAGGTCCAGCGCTGACACCTTAAACTG 2021
468 ATGAGCTGTGTTCAAGCTGTTTCCAGAGCTAGAGGTGAAGACCTTGGAAAGACACTA 527
2022 TGGAGGACAGCTGTGATTTCTTCAA--CACATGGAAGAGTCCCAACCCAGCGCTCA 2078
528 TCGTGCCCATGCTGTGATTTCTTGAATGAGATCTCTCAAGAGTGCAGACATGAGCACTGA 587
2079 CCCAGAGTTTGGGCACTTGTCCCGCGGCAAAACATGCTGTGTGTGGCACTGTGTGAAG 2138
588 CGGTGGAGTGGGCAAGTCAATGACGCAAAAGAGTCCCTGATTTCTCATCTGAGAGAC 647
2139 GAAAAGCCGACGCTCCGCGGAGTGTGGAAGGCCAGCTGCTCTTCCAGAGTT 2198
648 ACAAGGCTTTGCTTATACAAAGCTATAGAGGAGTGTGATACATGTGAGAGTGT 707
2199 CCATCTCTGTAGTGAACAACATGCAACATCATCTGTGTGATGAGAGCAGATATCAAGC 2258
708 CCGCTTTCAGAGACCGCGAGACAGTCTTTGTGTGATGATGAAGCCCATTTGGAAC 767

```

```

QY      2259 TG 2260
DB      768 TG 769

```

```

RESULT 5
BP705177
LOCUS      745 bp mRNA linear EST 19-JUL-2004
DEFINITION BP705177 Osada Taira anterior neuroectoderm (ANE) pcs105 cDNA
ACCESION   BP705177
VERSION    BP705177.1 GI:46053576
KEYWORDS   EST.
SOURCE      Xenopus laevis (African clawed frog)
ORGANISM   Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.

```

```

REFERENCE 1 (bases 1 to 745)
AUTHORS   Osada,S., Kitayama,A., Ueno,N. and Taira,M.
TITLE     Expression analysis of genes which are expressed in the anterior
          neuroectoderm of Xenopus embryos
JOURNAL    Unpublished (2004)
CONTACT    Contact: Masanori Taira

```

```

Department of Biological Sciences
Graduate School of Science, University of Tokyo; CREST, Japan
Science and Technology Corporation, Japan
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
Tel: 81-03-5841-4434
Fax: 81-03-5841-4434
Email: m.taira@biol.s.u-tokyo.ac.jp
URL: http://www.shigen.nig.ac.jp/hdrp/xenopus/est/.

```

FEATURES

SOURCE

```

1..745
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XU514b2lex"
/tissue_type="anterior neuroectoderm"
/dev_stage="late gastrula (stage 12.5)"
/clone_lib="Osada Taira anterior neuroectoderm (ANE)
pcs105 cDNA library"

```

ORIGIN

```

Query Match      1.9%; Score 113.4; DB 5; Length 745;
Best Local Similarity 52.2%; Pred. No. 4.5e-21;
Matches 325; Conservative 0; Mismatches 291; Indels 7; Gaps 3;

```

```

1600 TCCATATGCAACCTTCGCCCAACAAGGTGGAACCTTGGGGCTTGCACAGAGATCCTCAG 1659
120 TCCATATATCAACCGCTGATTAATCTTTACTTGGGGCTTGCACAGAGATCCTCAG 179
1660 CACTGAGTACTTACCAAGAGCTCATTCGATGTAAGAGCTGGGGAAGTGTCTTCAAGA 1719
180 CGTAGAGTGTACAAGAACTAATCGAGTACCAATTAAGAGAGTGTCTCTTCAAGT 239
1720 ACTGCAAGGCACTTCTTGTGATGTAATGAGTGGAGCTTAACCGTGCAGATGAAGAGCT 1779
240 ATGTGAAGAACTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 229
1780 ACTTTAAACCATTCGCAAGAGTTCACGACCAATCGACATGCTGTGATGAGAGAGTCT 1839
300 ATCATTCCTTATGTGGAACAACCTCTTTAAGCACTTGACATGACATGAGGTGAGATGCC 359
1840 ACAGCCAGATGTGTCAAAACCTGATCATTCGAAGAGCTGCGAGATATGAGGCAAGA 1899
360 ACATCTTGAGTGAAGAAAGCCACAGACTTTCAGGAGAGTGAACCTTTTGAAGATAAGA 419
1900 T--CGCTGAGAAATCCGTTGAAGTCAAAATCTTGGCATGCG-CGGAACCGGCATTCG 1955
420 TTCGGGCTGCTGTGGCATTAAGCTTTGTGTGAGAGGTATCGCTGTGATGATCATAG 479

```

QY 1956 CTTTCAATGAACCATCATCTTCTCTGTCAGACTGCAAAAGTCCAGCGCTCCACCTTA 2015
 Db 480 CCTTCATATGAGCCAGGGTCAAGTCTGTCTCCAGAGCCCGAGTGAATAATTAAGCATG 539
 QY 2016 AAAGTGTGAGAGCAACGCTGCAATCTT---CAACACCATTCGAAGAGTCCCAACCCAG 2072
 Db 540 ATACTATTTCTGGCATGACGAGTTCTTTGACCGCAACTCTTCGAAGTCCCAACCATG 599
 QY 2073 CCGTCAACCAAGGTTGGGCACTTTGTCCCGCGCGCAAAACATGTTGTGGCAACTG 2132
 Db 600 CCTGCTCTGTGGTGTGGGCACTGTATGACATCCAAAGAGTCATGATTTCTCATCATG 659
 QY 2133 GTGAAGAGAAAGCCGACGTCATCCGGGGAACGTGTGAAGCCCAAGTCTGTTCTTGGC 2192
 Db 660 GAGCCCATTAAGCTTTGCTTGTATTAAGGCTATTGAAGAGGTGAATCATATGTGGA 719
 QY 2193 CAGTTTCATCTCTGTATGATGAC 2215
 Db 720 CAGTTCTGCTTCCAGCAGCAC 742
 RESULT 6
 LOCUS CK807155 732 bp mRNA linear EST 26-FEB-2004
 DEFINITION AGENCOURT 19145816 NICHX_XGC_Te2 Xenopus laevis cDNA clone
 IMAGE:7208717 5', mRNA sequence.
 ACCESSION CK807155
 VERSION CK807155.1 GI:43396531
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Xenopus.
 1 (bases 1 to 732)
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Daniela S. Gerhard, Ph.D.
 COMMENT Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: Igor B. David
 cDNA Library Preparation: Express Genomics
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
 http://image.llnl.gov
 Plate: LHAM15095 row: a column: 03
 High quality sequence stop: 707.
 Location/Qualifiers
 1..732
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:7208717"
 /tissue_type="Pooled samples from 6 adult Xenopus testis"
 /lab_host="PH10B Tona"
 /clone_1lb="NICHX_XGC_Te2"
 /note="Organ: testis; Vector: pExpress-1; Site_1: EcorV;
 Site_2: NotI; RNA obtained from 6 adult male testis. cDNA
 was primed using oligo-dT primer:
 5'-pGACTAGTCTAGATCGCAGCGCGCGCC(T)25-3' and cloned into
 the EcorV/NotI sites of pExpress-1. Size-selection >1kb
 resulted in an average insert size of 1.25 kb. This is a
 primary library (normalized primary library is
 NICHX_XGC_Te2N) and was constructed by Express Genomics
 (Frederick, MD). Note: this is an NIH_XGC library"
 ORIGIN
 Query Match 1.9%; Score 112; DB 7; Length 732;

Best Local Similarity 54.0%; Pred. No. 1,2e-20;
 Matches 296; Conservative 0; Mismatches 245; Indels 7; Gaps 3;
 QY 1611 CCTTCGCGCAAGAGGTGGAACCTTGGGCTTGGCAACAGATCTTCACACTGAGTACC 1670
 Db 135 CCGACCGCTGATAGTACTTTTCTTTGGGCTTCCGACAGGAAGCACCCCTTAGATGC 194
 QY 1671 TACCAAGACTCATTCGATGATGATGAGTGGGGAAGTGTCAATTCAAGATGCAAGCA 1730
 Db 195 TACAAAACCTAATTGAGTACCATTAATAATGAGATCTGTCCCTCAAGATGTGAATAA 254
 QY 1731 TTTCTTTGGATGATAGTACGTGGGACTAACCCGTCAGATGAAGAACGCTACTTTAAAC 1790
 Db 255 TTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 314
 QY 1791 ATTGCAAGAGTCACTGACCATGACCATGACATGATGATGATGATGATGATGATGAT 1850
 Db 315 ATGTGGAACAATTCTTCAAGCATTGACATCAAGCTGAGAGCCCACTTTGAT 374
 QY 1851 GGTGCAAACTGATTCATACGACAGCTGACAGATGATGACGCAAAAGT---CGCTGCA 1907
 Db 375 GGAAGCGTGAAGACCTTCAGGCGCAGATGTGACCTGTTGAAGAAAGATTGCGGCTGCT 434
 QY 1908 GAATCCGTTGAAGTCAAAATCCTTGGCATCGG-CCGAAACGGCACATGCTTCAATTGA 1966
 Db 435 GGAAGGATTAACCTGTTGTTGAGAGTATCGGCTCTGATGGCCACATGCTTTCATGAG 494
 QY 1967 CCATCATCTTCTCTGTCAGACTGACCAAGATGTCAGCGCTCACCTTAAACTGTGAG 2026
 Db 495 CCAAGGTCAAGTCTGTGCTCCAGAGCCGAGTGAAGAAAGTGGCCATGAGCATATTCTG 554
 QY 2027 GACAAAGCTGATTTCTTCAA---CACCATGAAAGTTCACCAACCGCTCACCCAG 2083
 Db 555 GCTAATGACGCTTCTTGTGATGACCAACCTTCACCAAGTCCCAATGCTGACTGTG 614
 QY 2084 GGTGTGGGACCTTGTCCCGCGCGCAAAACATCGTGTGGCAACTGTGAAGGAAGA 2143
 Db 615 GGTGTGGGACCTTGTATGAGATTCGAAGAGTCAATGATTCATCATCGGAGCCCATTA 674
 QY 2144 GCCGAGCG 2151
 Db 675 GCCTTTCG 682
 RESULT 7
 LOCUS CK805715 760 bp mRNA linear EST 26-FEB-2004
 DEFINITION AGENCOURT 19145650 NICHX_XGC_Te2 Xenopus laevis cDNA clone
 IMAGE:7210866 5', mRNA sequence.
 ACCESSION CK805715
 VERSION CK805715.1 GI:43392452
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Xenopus.
 1 (bases 1 to 760)
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Daniela S. Gerhard, Ph.D.
 COMMENT Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: Igor B. David
 cDNA Library Preparation: Express Genomics
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
 http://image.llnl.gov

Plate: L1AM15100 row: j column: 16
High quality sequence stop: 667.
Location/Qualifiers

FEATURES

source

1. 760
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:9355"
/clone="IMAGE:7210866"
/tissue_type="Pooled samples from 6 adult Xenopus testis"
/lab_host="DH10B Tona"
/clone_1ib="NICHD XGC Te2"
/note="Torgan: testis; Vector: pExpress-1; Site_1: Scovy; Site_2: NotI; RNA obtained from 6 adult male testis. cDNA was primed using oligo-dT primer: 5'-pACTAGTCTAGATCGCGAGCGGCCCT(1)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >1kb resulted in an average insert size of 1.25 kb. This is a primary library (normalized primary library is NICHD_XGC_Te2N) and was constructed by Express Genomics (Frederick, MD). Note: this is an NIH_XGC library"

ORIGIN

Query Match 1.9%; Score 111.8; DB 7; Length 760;
Best Local Similarity 52.4%; Pred. No. 1.4e-20;
Matches 318; Conservative 0; Mismatches 282; Indels 7; Gaps 3;

1600 TCCATATGCGACCCCTTGGCAACAGGCTGGAACCTTGGGGCTTGGCAACAGATCCTCAC 1659
88 TCCATATTAATCCAAACCGCTGATAAATACCTTACTTGGGGCTTCCACAGAGACACGCG 147
1660 CACTAGATCTACCAAGAGCTATTCGATATGAAGCTGGGGAAGTCAATCCAGA 1719
148 CGCTAGATGCTACCAAGAGCTATTCGATATGAAGCTGGGGAAGTCAATCCAGA 207
1720 ACTGCAAGCATCTTCTGTGATGATATGATGATGATGATGATGATGATGATGATGAT 1779
208 ATGTGAAGATCTTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 267
1780 ACTTTAAACATTCGCAAGAGCTTCACTGACCAATGCAATGCTGTGATGATGATGATGAT 1839
268 ATCATTCCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 327
1840 AAGACCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1899
328 AACTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387
1900 T---CGCTCAGATCCGTTGAAGTCAATCTTGGGATCGG-CGGAACCGGACATCG 1955
388 TTCGGGCTGCTGGTGGCATTTGAACCTGTTGTTGAGAGTATCGGTCCTGATGATGATGAT 447
1956 CTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2015
448 CCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 507
2016 AACTGTGAGAGCAAGCTGATCTT---CAACACATGCAAGAGTCCCAACCCAG 2072
508 ATACTATTCGCGCAATGACAGCTTCTTTGACGGCACTCTCCAAAGTGCCAAACATG 567
2073 CCGTCAACCAAGGTTGGGCACTTGTCCCGCGGCAAAACATGCTGTGGTGAGCACTG 2132
568 CCTGACTGTGGGTGGGAGCTGTTATGATGATCCAAAGAGTCATGATTCATCATCAG 627
2133 GTGAAGAAAGAGGCAAGCCATCCGGGGAACGTGGAAGCCGAGGATGATGATGATGATG 2192
628 GAACCCATTAAGGCTTTGCTTTATTAAGGCTATTAAGAGGATGATGATGATGATGATG 687
QY 2193 CAGGTT 2199
DB 688 CAGTTTC 694

RESULT 8
BX714658

LOCUS BX714658 899 bp mRNA linear EST 18-NOV-2003
DEFINITION BX714658 XGC-tadpole Xenopus tropicalis cDNA clone TTPA01516 5', mRNA sequence.
ACCESSION BX714658
VERSION BX714658
KEYWORDS BX714658.1 GI:38386145
SOURCE EST.

ORGANISM

Xenopus tropicalis (western clawed frog)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;

REFERENCE

1 (bases 1 to 899)
Croning M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.
Unpublished (2003)

TITLE

JOURNAL

COMMENT

Sanger, Xenopus tropicalis EST project 2001 (11_2003)
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE ID: TTPA01516.p1kSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
cDNA was oligo dt primed from 5ug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli DH10B.

FEATURES

source

1. 899
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TTPA01516"
/dev_stage="tadpole (stage 35-40)"
/lab_host="E. coli DH10B"
/clone_1ib="XGC-tadpole"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dt primed from 5ug of poly A+ RNA from tadpole embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

Query Match 1.8%; Score 105.4; DB 5; Length 899;
Best Local Similarity 49.9%; Pred. No. 1.1e-18;
Matches 322; Conservative 0; Mismatches 316; Indels 7; Gaps 2;

1611 CCTTCGGCAACAGGCTGGAACCTTGGGCTTGGCAACGATCCTCACCTGATACC 1670
111 CCAACCGCTGATTAATACCTTAATTAGGGCTCCCAACAGAGACCCCACTAGATGC 170
1671 TACCAAGCTCATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1730
171 TACAAAACCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 230
1721 TCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1790
231 TTCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 290
1791 ATTCGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1850
291 ATGTGAACAACCTTCTTAAGCAATGACATGACAGCTGAGAAAGCCACATCTTGAT 350
1851 GGTGAAACCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1910
351 GGAATGCGACAGACCTTCAGGCAAGTGTACCTGTTTGAAGAAATTCGGCTCCT 410
1911 TCGTTGAAGTTCAATCCTTGCA---TCGGCGAAACGGCACATGCTTTCAATGA 1966
411 GGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 470
QY 1967 CCATATCTTCTGTGACAGACTGCAAAAGTCCAGGCGCTGACCCCTTAATGATGATG 2026

Db 471 CCAAGGCTCAAGTGTGCTTCACAGACCCGAGTAAACATATGGCCATGATACATTCTG 530
 QY 2027 GACACGCTGATGATTCTT---CAACACCATGAAAGAGTCCCAACCAAGCCGCTACCCAG 2083
 Db 531 GCCAATGACAGGTTCTTTGACGCGCAACCTCTCCAAAGTGCCACGATGCGCTGACTGTA 590
 QY 2084 GGTTTGGGCACTTTTGTCCCGCGGCAAAACATGTGTGTGGCACTGTGTGAAGAAAA 2143
 Db 591 GGTGTGGGCACTGTGTGATGACCTCAAGAAGATCATATTTCTATCACTGAGCCCATATA 650
 QY 2144 GCCGACGCGATCCGCGGAACTGTGAAGGCCGAGTACTCTTGTGCGCAGGTTCCATC 2203
 Db 651 GCATTTGCTTTGATTAAGGCAATGAGAGAGTGTAAATCATATGTGACACAGTTTCTGCT 710
 QY 2204 CTGTAGATGACACACATGCCACCATCATCTGTGATGAAGACGCA 2248
 Db 711 TTTCACACAGACCCACAGCACTGTGTGTTGCTGTGATGAGATGCA 755

RESULT 9 811 bp mRNA linear EST 18-AUG-2003
 CF346432
 LOCUS
 DEFINITION AGENCOURT 15227262 NICHD XGC Swb1n Xenopus tropicalis cDNA clone
 IMAGE:6998096 5', mRNA sequence.

ACCESSION CF346432 GI:33788300
 VERSION
 KEYWORDS
 SOURCE EST.
 ORGANISM Xenopus tropicalis (western clawed frog)
 Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Silurana.

REFERENCE
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Rob Granger, University of Virginia
 cDNA Library Preparation: Open Biosystems
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINT at:
 http://image.jnl.gov

FEATURES
 source
 Plate: LHM14686 row: a column: 07
 High quality sequence start: 15
 High quality sequence stop: 717.
 Location/Qualifiers
 1..811

/organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="IMAGE:6998096"
 /issue_type="whole body"
 /note="Vector: pEXpress-1. Site_1: EcoRV; Site_2: NotI;
 Bulk tissue was collected from a whole 10 month old male
 from the P6 strain. 1st strand cDNA was primed with a Not
 I - oligo (dT) primer, double-stranded cDNA was cloned into
 the Not I and EcoRV sites of pEXpress-1. Library was
 size-selected for >1.5 kb fragments for an average insert
 size of 1.92 kb. Library was normalized to Cots with a
 180-fold reduction of actin. A non-normalized version of
 this library is also available (NICHD XGC Swb1). Library
 was constructed by Open Biosystems (Huntsville, AL).
 PLEASE NOTE: This library contains high level of
 contamination by worm"

ORIGIN

Query Match 1.7%; Score 104.4; DB 7; Length 811;
 Best Local Similarity 50.5%; Pred. No. 2,2e-18;
 Matches 309; Conservative 0; Mismatches 296; Indels 7; Gaps 2;
 QY 1611 CCCTTGCCCAACAGAGGAGGAAACCTTGGGGCTTGCAACAGATCTCTACCACTAGTACC 1670
 Db 149 CCACCGCTGATTAATACCTTTACTTTAGGGCTCCCAACAGAAAGCAACCCACTAGATGC 208
 QY 1671 TACCAAGACCTCATTCGATGATGAAAGCTGGGAAAGTGTCACTTCAAGAACTGCAAGCA 1730
 Db 209 TACAAAAAATAATAGATACATATAAAACGAGATCTCTCTTCAAGATGTGAAAC 268
 QY 1731 TTCTTGTGATGAATATACGTGGGACTTAACCCGTGAGATGAAAAACAGCTACTTAAAC 1790
 Db 269 TTCAATATGATGATGATGTGGCTCTCCAGAGACACCTGAGAGCTATCATTCCTTC 328
 QY 1791 ATTGCAAGAGTTCATCTGACCATGACATGCTGTGATGAAGAGTCTACAGCCAGAT 1850
 Db 329 ATGTGAAACAATCTTTTAAGCATGACATGACATGAGGCTGAGAAATGCCACATCTTGAT 388
 QY 1851 GGTGCAAAACCTGATTCATACGAAGCACTGAGATGAGGCAAAAGATCGTCGAGA 1910
 Db 389 GGAATGACACAGACTTCAGGCAAGGTGTGACCTGTTGAAGAAAAATTGGGCTGCT 448
 QY 1911 TCCGTTGAAGTTCAATCCTTGGA---TGCGCGAAACGGCACTCGCTTCAATTGA 1966
 Db 449 GGTGGATGACCTGTTGTGTGAGGTATGTGCTGATGGCCACATACCTTTATATGAG 508
 QY 1967 CCATCATCTTCTCTGTCAGAGCTGACCAAGATTCAGAGCCCTGACCCCTAAACTGTGAG 2026
 Db 509 CCAGGCTCAAGTCTGTGCTTCAGAGACCCGAGTAAACATTTGCCATGATACATTCTG 568
 QY 2027 GACACGCTGATGATTCTT---CAACACCATGAAAGAGTCCCAACCAAGCCGCTACCCAG 2083
 Db 569 GCCAATGACAGGTTCTTTGACGCGCAACCTCTCCAAAGTGCCCAAGATGCGCTGACTGTA 628
 QY 2084 GGTTTGGGCACTTTGTCCCGCGGCAAAACATGTGTGTGGCACTGTGTGAAGAAAA 2143
 Db 629 GGTGTGGGCACTGTGTGATGACCTCAAGAAGATCATATTTCTATCACTGAGCCCATATA 668
 QY 2144 GCCGACGCGATCCGCGGAACTGTGAAGGCCGAGTACTGCTTGTGCCAGGTTCCATC 2203
 Db 689 GCATTTGCTTTGATTAAGCAATGAGAAAGTGTAAATCATATGTGACAGTTTCTGCT 748
 QY 2204 CTGTAGATGAC 2215
 Db 749 TTTCACACAGCAC 760

RESULT 10 896 bp mRNA linear EST 26-SEP-2003
 CF592490
 LOCUS
 DEFINITION AGENCOURT 15680772 NICHD XGC Swb1n Xenopus tropicalis cDNA clone
 IMAGE:7024229 5', mRNA sequence.
 ACCESSION CF592490 GI:36345132
 VERSION
 KEYWORDS
 SOURCE EST.
 ORGANISM Xenopus tropicalis (western clawed frog)
 Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus; Silurana.
 REFERENCE
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Rob Granger, University of Virginia

CDNA Library Preparation: Open Biosystems
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: L1M14753 row: f column: 03
 High quality sequence stop: 708.
 Location/Qualifiers

FEATURES

source

```
1..896
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:7024229"
/cisue_type="whole body"
/clone_lib="NICHU XGC SBDIN"
/note="Vector: pEXpress-1; Site 1: EcoRV; Site 2: NotI; Bulk tissue was collected from a whole 10 month old male from the P6 strain. 1st strand cDNA was primed with a Not I - oligo(dT) primer, double-stranded cDNA was cloned into the Not I and EcoRV sites of pEXpress-1. Library was size-selected for >1.5 kb fragments for an average insert size of 1.92 kb. Library was normalized to Cots with a 180-fold reduction of actin. A non-normalized version of this library is also available (NICHU XGC SBD1). Library was constructed by Open Biosystems (Huntsville, AL). PLEASE NOTE: This library contains high level of contamination by worm"
```

ORIGIN

Query Match 1.7%; Score 104.4; DB 7; Length 896;
 Best Local Similarity 50.5%; Pred. No. 2.3e-18;

Matches 309; Conservative 0; Mismatches 296; Indels 7; Gaps 2;

```
1611 CCTTCGCCAAGAGGTGAACTTGGGCTTGGCAAGATCTCACCAGTACG 1670
143 CCACCGCTGATTAATCTTATAGGGCTCCCAAGAGCAAGCCACCTAGATGC 202
1671 TACCAAGAGCTCAATCGCATGTATGAGCTGGGGAAGTCTCAAGAACTGCAAGGA 1730
203 TACAAAACCTAATAGATACCAATAAAACGAGATCTCTCTCAAGTATGTGAAACA 262
1731 TTCTTGTTGATGAATAGTGGGACTAACCCGTAAGATGAAACAGCTACTTAAAC 1790
263 TTCAATATGATGATGATATGAGCTTCCAGGAGACCTCGAGAGCTATATCTTC 322
1791 ATTGGCAAGATTCATGACCAATCGACATCGTTGATGAAGAGTCTACAGCCAGAT 1850
323 ATGTGGAACAATCTTTAAGCAATCGACATCGAGGTGAGAAATGCCACATCTTGAT 382
1851 GGTGCAAAACCTGATTCATACGAGAGTGCAGAGTATGAGGCAAAAGTGGCTGCAAA 1910
383 GGAAATGCCACAGACCTTCAGCAGAGTGTGACTGTGTTGAAGAAAGATCGGGCTGT 442
1911 TCCGTTGAAGTCAAAATCTTGCA----TCGGCGGAAACGGCACTGCTTCAATTGA 1966
443 GGTGGGATGAGCTGTTGTTGAGGATATGTTGCTCGATGCGCACATGACTTTATATG 502
1967 CCATCATCTTCTCTGTCAGAGCTGACAAAGTCCAGCGCTGACCCCTTAAACTGTGAG 2026
503 CCAAGGTCAAGTCTGTGCTCCAGGACCGAGTAAACATTTGGCCATGATACATTTG 562
2027 GACAAAGCTCGATTTCT---CAACACATCGAAAGAGTCCCAAGCCGCTCACCCG 2083
563 GCAATGACAGGTCTTTGACGCGCACTCTCCAAAGTCCAAAGATGCGCTGACTGTA 622
2084 GGTGGGCACTTTGCTCCGCGCGCAAAACATGTTGGTGGCAATGATGAAGGAAA 2143
623 GGTGAGGAGCTGTATGATCTCCAAAGATCATATCTATATCACTGAGACCATATA 682
2144 GCCGAGCCATCCGCGGAACTGTGAGAGCCCAAGTACGCTTCTTGGCCAGATTTCATC 2203
683 GCATTTGCTTTATAGGCAATTGAGGAAGTGTAAATCATATGTGACAGATTCTGTG 742
```

QY 2204 CTGTAGATGCAC 2215
 Db 743 TTTTCACAGCAC 754

RESULT 11

CR696680 1177 bp mRNA linear HTC 19-AUG-2004

LOCUS Tetraodon nigroviridis full-length cDNA.
 DEFINITION CR696680
 VERSION CR696680.1 GI:51194589
 KEYWORDS HTC; cDNA; full-length; Tetraodon nigroviridis.
 SOURCE Tetraodon nigroviridis
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE

AUTHORS
 TITLE
 JOURNAL

COMMENT

The sequences are based on single pass reads.
 More information available at
<http://www.genoscope.cns.fr/tetraodon>.

FEATURES

source

```
1..1177
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/cisue_type="5yes"
```

ORIGIN

Query Match 1.7%; Score 104.4; DB 3; Length 1177;
 Best Local Similarity 49.7%; Pred. No. 2.5e-18;

Matches 324; Conservative 0; Mismatches 321; Indels 7; Gaps 2;

```
1603 TAATGCAACCTTCGCCAAGAGGTGAACTTGGGCTTGGCAAGATCTTACAC 1662
336 TATTCATCCCGCGCGGACCGGTATTTTACCTCGGGCTGCCACAGAGCACTCTC 395
1663 TGAGTACCTTACCAAGAGTCTATGCAATGAGTGGGGAATGTCATTCAGAACT 1722
396 TGGTTGCTTACAAAGAGTGAATGATTAACAAAGATGAGAAAGTCTGTTGGTACG 455
1723 GCAAGCAATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1782
456 TGAAGACCTTCAACATGAGCAATTCGTTGCTGCCAGGAGCAAGCCGAGAGTACC 515
1783 TTAACCAATTCGCAAGAGTTCATGACCAATGACATGATGATGATGATGATGAT 1842
516 ACTCTTCATGAGGAGCAACTCTTCAAGACATGATGATGATGATGATGATGATGAT 575
1843 GCCCAGATGAGCAACCTGATTCATACGAGAGCTGCAAGATGATGATGATGATGAT 1902
576 TTTTGAATGGAACCGCCGCACTGCAAGAGAGTGGAAAGCTTGCAGAGAGAAATG 635
1903 CTGAGAAATCCGTTGAAGTTCAATCTTGCA----TCGGCGGAAACGGCACTGCTT 1958
636 TGGAGGCTGAGGATTCAGCTGTTTGTGCGAAGAAATGAGACGGACCAATTTGCT 695
1959 TCATTTGAACATCATCTTCTCTGTCAGAGATGACAAAGTCCAGGCGTGCACCTTAAA 2018
696 TCAAGAGCGCGGTTCCAGTTTGTGTCAGAAACCGGATGAAAGACCTGCGCAAGACA 755
2019 CTGTGAGAGCAAGCTGATTTCTTCAAC---ACATGGAAGAGTCCCAAGCCAGCGCG 2075
756 CCATATTGCAACCGCGGCTTCTTGAAGGAGACTTTCCAAAATGCTCCACCAAGGCC 815
2076 TCACCAAGGTTTGGGCACTTTGTCGCCGCGCAAAACATGATGATGATGATGATGAT 2135
```


Db 816 TGACTGTGGCGCGTGGGCACTGTATGAGCGCCAAAGAGTCATGATTTCTATACAGCGGCG 875
Qy 2136 AAGAAAAGCCGACCGCATCCGCGGAACTGTGAAAGGCCCATGACGTCTTGTGCCAG 2195
Db 876 CACACAAGGCTCTGCGCTTGTGGCCAAAGCTATCGAGGAGGCGGTGAATCATATGTGACCG 935
Qy 2196 GTTCCATCCTGTAGATGACAAATGCGACCATCTGTTGATGTAAGCAGC 2247
Db 936 TCTGTGCTTCCAGACACCCACAGTTCATCTTGTGTGCGACGAGACGC 987

RESULT 12
CR690088 1178 bp mRNA linear HTC 19-AUG-2004
LOCUS Tetraodon nigroviridis full-length cDNA.
DEFINITION CR690088
ACCESSION CR690088.1 GI:51187995
VERSION HTC; cDNA; full-length; Tetraodon nigroviridis.
KEYWORDS Tetraodon nigroviridis
SOURCE Tetraodon nigroviridis
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1178)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage - 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
source 1..1178
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/cissue_type="Eyes"

ORIGIN
Query Match 1.7%; Score 104.4; DB 3; Length 1178;
Best Local Similarity 49.7%; Pred. No. 2.5e-18;
Matches 324; Conservative 0; Mismatches 321; Indels 7; Gaps 2;
Qy 1603 TAATCGACCCCTTGGCAACAGGCTGAACTTGGGCGCTTGACAGAGATCTTCAACAC 1662
Db 303 TATTCAATCCCGGCGCGGACCGGTATTTTACCTGGGGCTGCCACAGAGACATCTCTC 362
Qy 1663 TGAATACCTTACCAAGAGCTCATTCGATGTATGAACTGGGGAAGTGTCTTCAAGAACT 1722
Db 363 TGGGTGTGTAACAAGAGCTGATGAGTATTAACAAGATGAGAAGTCTGTTCACTAGC 422
Qy 1723 GCAAGGATTTCTTTGGATGATAGTGGGACTAACCCGTGACATGAAAAAGCTACT 1782
Db 423 TGAAGCTTTCAACATGAGCAATATGTGTCTCCCGAGAGCCACCCCGAGAGCTAC 482
Qy 1783 TTTAAACATTCGCAAGAGTTCATGACCATGACATCGTTGATGAAGAGTCTACA 1842
Db 483 ACTCCCTTACGTGGAGCACTTCTTGAAGACATGACATCAAGCGGAAAAACATCACA 542
Qy 1843 GCCCAGATGTGCAAACTCTGATCATACGAGACAGCTGACAGTATGAGGCAAAAGATCG 1902
Db 543 TTTTACAGCGCAACGCGCGGACCTGACAGAGAGTGCAGAGCTTTCGAGAGAAATGG 602
Qy 1903 CTGCAATCCGTTGAAGTTCAATCCTTGGCA----TGGCGGAAACGGCAATGCTT 1958
Db 603 TGGAGCTGTGTGATTCACGCTTGTGTGCGAGGAGTTGGACCGGACGGCAATTTGCT 662
Qy 1959 TCATTGAACCATCATCTTCTGTCTGTCAGACTGACAAAGTCCAGGCGCTGCACCTTAAA 2018
Db 663 TCAAGAGCGCGGTTCCAGTTGTGTGTCAGAACGCGGGGTGAAGAACCTTGGCCAAAGACA 722

Qy 2019 CTGTGAGAGACAACGCTGATTTCTTAAC--ACCATGAAAGAGTCCCAACCCAGCGC 2075
Db 723 CCATCATTTCCAAACGACGCTTCTTGACGGGAGCTTCCAAAGTCCCAACGAGCC 782
Qy 2076 TCACCCAGGATTTGGGCACTTGTCCCGCGCCAAACATCGTGTGTGTCGCAACTGTG 2135
Db 783 TGACTGTGGCGGTGGGCACTGTATGAGAGCCAAAGAGTCAATGATTTCTATTAACGGCG 842
Qy 2136 AAGAAAAGCCGACCGCATCCGCGGAACTGTGAAAGGCCCAAGTACTGCTTGTGCCAG 2195
Db 843 CACACAAGGCTCTGCGCTTGTGCGCAAGCTATCGAGAGGCGCTTATCATATGTGACCG 902
Qy 2196 GTTCCATCCTGTAGATGACAAATGCGACCATCTGTTGATGTAAGCAGC 2247
Db 903 TTTGTGCTTCCAGACACCCACAGTTCATCTTGTGTGCGACGAGACGC 954

RESULT 13
CR701935 1206 bp mRNA linear HTC 19-AUG-2004
LOCUS Tetraodon nigroviridis full-length cDNA.
DEFINITION CR701935
ACCESSION CR701935.1 GI:51199844
VERSION HTC; cDNA; full-length; Tetraodon nigroviridis.
KEYWORDS Tetraodon nigroviridis
SOURCE Tetraodon nigroviridis
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1206)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2004) Genoscope - Centre National de Sequencage - 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
(E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT The sequences are based on single pass reads.
More information available at
http://www.genoscope.cns.fr/tetraodon.
Location/Qualifiers
source 1..1206
/organism="Tetraodon nigroviridis"
/mol_type="mRNA"
/db_xref="taxon:99883"
/cissue_type="Eyes"

ORIGIN
Query Match 1.7%; Score 104.4; DB 3; Length 1206;
Best Local Similarity 49.7%; Pred. No. 2.6e-18;
Matches 324; Conservative 0; Mismatches 321; Indels 7; Gaps 2;
Qy 1603 TAATCGACCCCTTGGCAACAGGCTGAACTTGGGCGCTTGACAGAGATCTTCAACAC 1662
Db 348 TATTCAATCCCGGCGCGGACCGGTATTTTACCTGGGGCTGCCACAGAGACATCTCTC 407
Qy 1663 TGAATACCTTACCAAGAGCTCATTCGATGTATGAACTGGGGAAGTGTCTTCAAGAACT 1722
Db 408 TGGGTGTGTAACAAGAGCTGATGAGTATTAACAAGATGAGAAGTCTGTTCCGTAGC 467
Qy 1723 GCAAGGATTTCTTTGGATGATAGTGGGACTAACCCGTGACATGAAAAAGCTACT 1782
Db 468 TGAAGCTTTCAACATGAGCAATATGTGTCTCCCGAGAGCCACCCCGAGAGCTAC 527
Qy 1783 TTTAAACATTCGCAAGAGTTCATGACCATGACATCGTTGATGAAGAGTCTACA 1842
Db 528 ACTCCCTTACGTGGAGCACTTCTTCAAGCAATGACATCAAGCGGAAAAACATCACA 587
Qy 1843 GCCCAGATGTGCAAACTCTGATCATACGAGAGCTGACAGTATGAGGCAAAAGATCG 1902
Db 588 TTTTGAATGCAACCGCGCGGACCTGACAGAGAGTGCAGAGCTTTCGAGAGAAATGG 647
Qy 1903 CTGCAATCCGTTGAAGTTCAATCCTTGGCA----TGGCGGAAACGGCAATGCTT 1958

Db 648 TGGAGCTGGTGGATTCAGCTGTTGTGGAGGAATTGGACCGGACCATTTGCT 707
 Qy 1959 TCATTGAACCATCATCTTCTGTGCAGACTGACAAAGSTCCAGGCGTGCACCTTAAA 2018
 Db 708 TCACGAGCGCGGTTCCACTTTGGTGTCCAGAACCGCGGTGAAGACCTGGGCAAGACA 767
 Qy 2019 CTGTGAGGACAAACGCTCGATTCTTCAAC--ACATGGAAGAGGTCCCAACCCAGCGC 2075
 Db 768 CCATCATCTGCCAAGCGCGCTTCTTTGACGGGAGACCTCTCCAAAGGCCCAACGCGCC 827
 Qy 2076 TCACCGAGGGTTTGGGCACTTTGTCCCGGCGCAAAACATCGTGTGGCAATGCG 2135
 Db 828 TCATCTGTGGCGGTGGGCACTGTCAATGACGCCAAGAGGTCAATGATTTCAATCAGGCGC 887
 Qy 2136 AAGGAAAGCCGACCGCATCCGCGAATCTGTGAAGGCCCACTGCTTTGCCAG 2195
 Db 888 CACACAAAGCTCTGGCTTTGGCCAAAGCTATCGAAGAGGCGGTGAATCATGTGACCG 947
 Qy 2196 GTTCCATCTGTAGATGCACAAATGCCACCATCTGTTGATGAGCAGC 2247
 Db 948 TCTCTGCCCTTCAGACGACCCACCATCTCTGTGTGCGACGAGCAGC 999

RESULT 14
 LOCUS BM017859 818 bp mRNA linear EST 30-OCT-2001
 DEFINITION 60364535F1 NIH_MGC_98 Homo sapiens cDNA clone IMAGE:5427056 5',
 mRNA sequence.
 ACCESSION BM017859
 VERSION BM017859.1 GI:16532213
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE NIH-MGC <http://mgs.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cga@bbs-rcmail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLT)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNLT at:
<http://image.llnl.gov>
 Plate: L1CM1892 row: e column: 09
 High quality sequence stop: 803.
 Location/Qualifiers

FEATURES
 source 1..818
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5427056"
 /tissue_type="astrocytoma grade IV, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="NIH_MGC_98"
 /note="Organ: Brain; Vector: pOTB7; Site 1: XhoI; Site 2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Library constructed by Ling Hong
 in the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 1.7%; Score 103.8; DB 4; Length 818;
 Best Local Similarity 49.5%; Pred. No. 3.3e-18;
 Matches 327; Conservative 0; Mismatches 327; Indels 7; Gaps 2;

Qy 1610 ACCCTTCCCAACAAAGGTGAACTTTGGGCTTGCAACAGAGATCTTACCACTGAGTAC 1669
 Db 78 ACCCAAGGACAGAGAGTACTTACCTTGGGCTCCCACTGGAGTACCCCACTTGGCTG 137
 Qy 1670 CTACCAAGAGCTCATGTGCAATGTAAGCTGGGGAAGTGCATTCAGAACTGCAAGGC 1729
 Db 138 CTACAGAAAGCTGATTAATCTATTAAGATGGGAGCTGTCTTTAAATATGTAAGAC 197
 Qy 1730 ATTCTTGTGGAATGAATAGTGGGACTAACCCTGACATGAAACACACTTATTAAC 1789
 Db 198 CTTCAACATGATGATGATGATGAGGCTTCTTCAGAACACCGGAGATTTACACTCTCT 257
 Qy 1790 CATTGCAAAAGACTTCATGACACATGATGACATGTTATGAAGGCTTACAGCCCA 1849
 Db 258 CATGTGGAACAACTTCTTCAAGCAATGACATCCACCAAAACACCTCAATTGTGA 317
 Qy 1850 TGGTCAAAACCTGATTCATACGAAGCAGCTGACAGATGATGAGGCAAGATCGCTGAGA 1909
 Db 318 TGGGAATGACGTGACCTACAGGAGAAATGTGATGCTTTGAAGAGATTAAGGCTGC 377
 Qy 1910 ATCCGTTGAAGTTCAAAATCTTGGCATGGCGG---AAACGGACATGCTTTGATTGA 1965
 Db 378 AGGTGGATTCAGACTAATTGTTGGAGCGATCGGCTGATGAGACATTTGCCCTTCAAGA 437
 Qy 1966 ACCATCATCTTCTCTGTGACAGACTGACAAAGTCCAGGCGCTGCACCTTAAATGTGGA 2025
 Db 438 GCCAGGCTTCAGTCTGTGTGTCGACAGCCGCTGTGAAGCGCTGGCAGATGATTCATCT 497
 Qy 2026 GGACAAACGCTGATCTTCTCAA---CACCATGAAGAGGTCCCAACCCAGCGCTCACCCA 2082
 Db 498 GGCCAAATCTGATGTTCTTTCATGAGAACTCAACCAAGGTGCCACATGAGCGCTTGA 557
 Qy 2083 GGGTTGGACACTTGTGCTCCCGCGCAAAACATGCTGTTGTGTGCACTGTGTAAGAA 2142
 Db 558 GGGGGTGGGCACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 617
 Qy 2143 AGCCGACCCCATCCCGGGAACCTGTGAGAGCGCCAGTGTGCTGTTTCCCAAGTTTCAT 2202
 Db 618 GGCATTTCTCTGTACAAAGCCCATGAGAGGAGAGTGAACCAATGTGATCGCTGTGCG 677
 Qy 2203 CCTGTAGATGACAAACATGACCACTCATCTGTGATGAGAGCAGATGATCCAGCTGGA 2262
 Db 678 CTTCCAGACAGATTCGCCCGACCGGTTGTGTGTGATGACAGAGATGCCACTTGAGACTGA 737
 Qy 2263 A 2263
 Db 738 A 738

RESULT 15
 LOCUS AL656223 662 bp mRNA linear EST 07-NOV-2003
 DEFINITION AL656223 XGC-neurula Xenopus tropicalis cDNA clone tReu035007 5',
 mRNA sequence.
 ACCESSION AL656223
 VERSION AL656223.2 GI:38223958
 KEYWORDS EST.
 SOURCE Xenopus tropicalis (western clawed frog)
 ORGANISM Xenopus tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
 Xenopodine; Xenopus; Silurana.
 REFERENCE 1 (bases 1 to 662)
 AUTHORS Crotting,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
 TITLE Sanger Xenopus tropicalis EST project 2001 (11_2003)
 JOURNAL Unpublished (2003)
 COMMENT On Dec 13, 2001 this sequence version replaced gi:11768464.
 Contact: Huckle E
 Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Aaron M. Zorn.
cDNA was oligo dt primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.
Vector: pCS107; Site_1: EcoRI; Site_2: NotI
Host: Escherichia coli DH10B
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu035007.plksp6
Sequencing primer: SP6.
Location/Qualifiers

FEATURES

source

1..662

/organism="Xenopus tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="TNeu035007"

/dev_stage="neurula"

/lab_host="Escherichia coli DH10B"

/clone_1ib="XGC-neurula"

/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA

was oligo dt primed from 5ug of poly A+ RNA from neurula.

EcoRI-NotI cut cDNA was then ligated into pCS107 with

EcoRI at the 5' end and NotI at the 3' end."

ORIGIN

```
Query Match      1.7%; Score 103.2; DB 1; Length 662;
Best Local Similarity 52.0%; Pred. No. 4.5e-18;
Matches 282; Conservative 0; Mismatches 253; Indels 7; Gaps 2;

QY 1611 CCCCTGCGCAAGAGGTGAAACCTTGAGGCTTGCAACAGATCCTCACCAGTGAATACC 1670
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 111 CCAACCGCTGATTAATACCTTCTTAGGGCTCCCAAGAGAGACCCCACTAGATGC 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1671 TACCAAGAGCTCATTCGATGATGAAGCTGGGAAAGTGTCAITCAAGACTGCAAGCA 1730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 171 TACAATAAATTAATAGATACCAATAAAGAGATCTCTCTCAAGTATGTGAATAACA 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1731 TTCTGTGATGATAGTATAGTGGAGCTAACCGGAGCATGAATAACGTAATTAAACC 1790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 231 TTCAATATGATGATAGTATGAGCTTCCAGGAGACACCTGAGAGCTATCATTCCTTC 290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1791 ATTCGCAAGAGTCTCACTACCAATCGACATCGTTGATGAAGAGTCTACAGCCAGAT 1850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 291 ATGTGAACAATCTTCTTAAGCAATCGACATCAAGCGCTGAGATGCCACATCTTGAT 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1851 GGTGCAAACTGTATCATACGAAGCAGCTGCAAGATGAGGCAAGATCGCTGCAGAA 1910
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 351 GGAAATGCCACAGACCTTCAGGACAGTGTGACTGTGTAAGAAAAAGATCGGGCTGCT 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1911 TCCGTTGAAGTTCGAATCCTTGG----CATCGCGGAAACGGACATCGCTTTCATTGAA 1966
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 411 GGTGGATGAGCTGTTGTGAGGTATGTGCTCTGATGGCCACATAGCTTTTAATGAG 470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1967 CCATCATCTTCTCTGTCAGAGACTGACAAAGTTCAGAGCGCTGCACCCCTAAACTGTGAG 2026
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 471 CCAAGGTCAAGTCTGTCTCCAGAGACCCGATTAACATGTGCCATGATACATTCG 530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 2027 GACAAAGCTGATTTCTT---CAACACCATGAAAGAGTCCCAACCGCGCTCACCCAG 2083
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 531 GCCAATGCAAGGTTCTTGAACGCAACCTCCAAAGTGCACAAAGATGAGCCCTGACTGTA 590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 2084 GGTTCGGGACATTTGTCCCGCGCGCAAAACATGTGTTGTGTGCAACTGTGAAGAAAA 2143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 591 GGTGTGGGACTGTGTGATGATCCCAAGAGTCAATGATTCTAATCATCTGAGGCCATAAA 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 2144 GC 2145
    ||
DB 651 GC 652
```

Search completed: March 10, 2005, 08:24:38
Job time : 16585 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 21:56:27 ; Search time 168 Seconds
(without alignments)
1521.718 Million cell updates/sec

Title: US-10-019-284B-2

Perfect score: 3342
Sequence: 1 MDKDLAQRILRDIGEDNI.....IEAGNLINAKKEAVPATP 661

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp19808:*
2: geneseqp19908:*
3: geneseqp20008:*
4: geneseqp20018:*
5: geneseqp20028:*
6: geneseqp20038:*
7: geneseqp20039:*
8: geneseqp20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3342	100.0	661	4 AAB69080	Aab69080 Brevibact
2	3310	99.0	661	4 AAG92650	Aag92650 C glutam
3	2332	69.8	468	4 AAB66707	Aab66707 C.glutam
4	2332	69.8	468	5 ABG80325	Abg80325 C. glutam
5	1806	54.0	362	4 AAB66708	Aab66708 C.glutam
6	965	28.9	627	8 ADK46509	Adk46509 Streptoco
7	965	28.9	627	8 ADK94615	Adk94615 Novel S.
8	960.5	28.7	496	7 ADH87937	Adh87937 Enterococ
9	957	28.6	617	5 ABB847495	Abb847495 Listeria
10	957	28.6	617	6 ABU32951	Abu32951 Protein e
11	956	28.6	627	4 AAU37874	Aau37874 Streptoco
12	949	28.4	655	6 ABU02378	Abu02378 S. pneumo
13	949	28.4	655	6 ABP81458	Abp81458 Streptoco
14	935.5	28.0	620	5 ABP27216	Abp27216 Streptoco
15	935.5	28.0	655	8 ADK46443	Adk46443 Streptoco
16	931	27.9	676	5 ABP25654	Abp25654 Streptoco
17	929.5	27.8	478	6 ABM71603	Abm71603 Staphyloc
18	924.5	27.7	634	5 ABB849833	Abb849833 Listeria
19	923	27.6	628	7 ADC95661	Adc95661 E. faeciu
20	920	27.5	674	5 ABP25655	Abp25655 Streptoco
21	880.5	26.3	639	5 ABP27215	Abp27215 Streptoco
22	875.5	26.2	656	7 ABG61695	Abg61695 Klebsiell
23	859	25.7	617	5 ABB849923	Abb849923 Listeria
24	855	25.6	670	7 ADH86721	Adh86721 Enterococ
25	829	24.8	618	5 ABB47649	Abb47649 Listeria

ALIGNMENTS

26	826	24.7	631	7 ADC94275	Adc94275 E. faeciu
27	822.5	24.6	636	5 ABB54803	Abb54803 Lactococc
28	812	24.3	483	7 ABO66501	AbO66501 Klebsiell
29	802.5	24.0	634	7 ADP06734	Adp06734 Bacterial
30	801.5	24.0	612	6 ABU45919	Abu45919 Protein e
31	799.5	23.9	612	6 ABU00950	Abu00950 S. pneumo
32	799.5	23.9	612	6 ABP81336	Abp81336 Streptoco
33	796.5	23.8	612	8 ADK46396	Adk46396 Streptoco
34	791	23.7	622	5 ABR28755	AbR28755 Streptoco
35	785.5	23.5	620	5 ABR28854	AbR28854 Streptoco
36	785.5	23.5	620	6 ABU46515	Abu46515 Protein e
37	747.5	22.4	640	5 ABB49939	Abb49939 Listeria
38	738	22.1	379	6 ABU42285	Abu42285 Protein e
39	737	22.1	583	7 ADC97184	Adc97184 E. faeciu
40	728	21.8	618	8 ADK61668	Adk61668 Coryneb
41	725.5	21.7	683	4 AAG93207	Aag93207 C glutam
42	725.5	21.7	683	4 AAB66721	Aab66721 C.glutam
43	725.5	21.7	683	4 ABG80333	Abg80333 C. glutam
44	719	21.5	494	5 ABB49322	Abb49322 Listeria
45	691	20.7	704	4 AAU60962	Aau60962 Propionib

RESULT 1	ADK46509	standard; protein; 661 AA.
ID	AAB69080	
AC	AAB69080;	
DT	11-SEP-2003 (revised)	
DT	20-APR-2001 (first entry)	
DE	Brevibacterium lactofermentum sucrose PTS enzyme II protein SEQ ID NO:2.	
KW	Brevibacterium lactofermentum; sucrose PTS enzyme II; sucrose;	
KW	phosphoenolpyruvate:carboxylate phosphotransferase system; glucose;	
KW	coryneform bacterium; phosphoenolpyruvate-sugar transport system.	
OS	Corynebacterium glutamicum.	
PN	MO200102584-AL.	
PD	11-JAN-2001.	
PF	30-JUN-2000; 2000MO-JP004348.	
PR	02-JUL-1999; 99JP-00189512.	
PA	(AJIN) AJINOMOTO CO INC.	
PI	Izui M, Sugimoto M, Nakamatsu T, Kurahashi O;	
DR	WPI; 2001-138150/14.	
DR	N-PSDB; AAF32543.	
PT	Gene encoding sucrose phosphoenolpyruvate-sugar transport system enzyme	
PT	II obtained by cassette ligation-mediated amplification of downstream	
PT	domain of coryneform bacterium sucrose gene, with sucrose-binding	
PT	activity.	
XX	Claim 1; Page 29-32; 45pp; Japanese.	
XX	The present sequence represents the Brevibacterium lactofermentum sucrose	
XX	PTS (phosphoenolpyruvate:carboxylate phosphotransferase system or	
XX	phosphoenolpyruvate-sugar transport system) enzyme II, which has sucrose-	
XX	binding activity. A coryneform bacteria produced with the sucrose PTS	
XX	enzyme II gene can have more efficient sugar uptake, and improved auto-	
XX	acid and nucleic acid productivity. The sucrose PTS gene and its	
XX	disrupted gene, such as one without the sucrose PTS function, can be used	
XX	to produce new breeds of coryneform bacterial strains to uptake sugar	
XX	more efficiently e.g. glucose only or and sucrose, and can have improved	

CC amino-acid and nucleic acid productivity. (Updated on 11-SEP-2003 to
CC standardise OS field)

XX Sequence 661 AA;

Query Match 100.0%; Score 3342; DB 4; Length 661;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 MDHKDLAQRILRDIGGEDNIVAAACATRLRLVLTQTKVDVRSQSLDDPDLKGTETGGM 60
DB      1 MDHKDLAQRILRDIGGEDNIVAAACATRLRLVLTQTKVDVRSQSLDDPDLKGTETGGM 60
QY      61 FOIIVGPGVDVHFPELDDATSKDIASVTEOLKDVANNANMFSRAVKVLADIPVPLPI 120
DB      61 FOIIVGPGVDVHFPELDDATSKDIASVTEOLKDVANNANMFSRAVKVLADIPVPLPI 120
QY      121 LVGGGLMAINNVLVAQDLFGPQSLVEMFPOISGVAEMINIMASAPFAFLPVLVGFTATK 180
DB      121 LVGGGLMAINNVLVAQDLFGPQSLVEMFPOISGVAEMINIMASAPFAFLPVLVGFTATK 180
QY      181 REGNEFLGAGIGMANVPEPTLVNGYDVAATMTAGEMPMSLFGLDVAQGYGTVLPLV 240
DB      181 REGNEFLGAGIGMANVPEPTLVNGYDVAATMTAGEMPMSLFGLDVAQGYGTVLPLV 240
QY      241 VSMILATIEKFLHKLMTADFLITPVLTLLTGFTITAIQPMRWGDLIAHGLQGLY 300
DB      241 VSMILATIEKFLHKLMTADFLITPVLTLLTGFTITAIQPMRWGDLIAHGLQGLY 300
QY      301 DFGPVGGLFGLVYSPVITGTHOSFPPIELFNOGGSFIFATSMANIOGAACLA 360
DB      301 DFGPVGGLFGLVYSPVITGTHOSFPPIELFNOGGSFIFATSMANIOGAACLA 360
QY      361 FFLASEKIKGLAGASGSAVLGITEPAIFGVNRLRMPFYIGITAIIGALLALPDIK 420
DB      361 FFLASEKIKGLAGASGSAVLGITEPAIFGVNRLRMPFYIGITAIIGALLALPDIK 420
QY      421 AVALGAAGFLGVSIDAPDMNMFVCAVVTPTIAGAAIAYGLVYVRNGSIDDPATAP 480
DB      421 AVALGAAGFLGVSIDAPDMNMFVCAVVTPTIAGAAIAYGLVYVRNGSIDDPATAP 480
QY      481 VPAGTTKAABAAPAESNDSTIIQAPLTGEAIALSSVDPAMFSGKLSGVAIVPTKGL 540
DB      481 VPAGTTKAABAAPAESNDSTIIQAPLTGEAIALSSVDPAMFSGKLSGVAIVPTKGL 540
QY      541 VSPVSGKIVVAPPSGHAFAVRTKAEDGSNVDILMHIGFDTVNLNGTHFNPLKKQDE 600
DB      541 VSPVSGKIVVAPPSGHAFAVRTKAEDGSNVDILMHIGFDTVNLNGTHFNPLKKQDE 600
QY      601 GELLCEPDIDAIKAAGYEVTTPIVVSNYKKTGPVNTYGLGEIEAGANLLNVAKKAV 660
DB      601 GELLCEPDIDAIKAAGYEVTTPIVVSNYKKTGPVNTYGLGEIEAGANLLNVAKKAV 660
QY      661 P 661
DB      661 P 661

```

RESULT 2
AAG92650
ID AAG92650 standard; protein; 661 AA.
XX
AC AAG92650;

XX 26-SEP-2001 (first entry)

DE C glutamicum protein fragment SEQ ID NO: 6404.

XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
KM organic acid synthesis.
XX
OS Corynebacterium glutamicum.

XX

PN EPI108790-A2.

XX 20-JUN-2001.

XX 18-DEC-2000; 2000EP-00127688.

XX 16-DEC-1999; 99JP-00377484.

XX 07-APR-2000; 2000JP-00159162.

XX 03-AUG-2000; 2000JP-00280988.

XX (KYOWA) KYOWA HAKKO KOGYO KK.

XX Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

XX Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX WPI; 2001-376931/40.

XX N-PSDB; AAH67869.

XX Novel polynucleotides derived from Corynebacterium bacteria, for identifying

XX mutation point of a gene, measuring expression of a gene, analyzing

XX expression profile or pattern of a gene and identifying homologous gene.

XX Claim 17; SEQ ID NO 6404; 246pp + Sequence Listing; English.

XX The present invention provides a number of nucleotide and protein

XX sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These

XX are useful for identifying the mutation point of a gene derived from a

XX mutant of corynebacterium bacterium, measuring expression amount and analysing

XX the expression profile or expression pattern of a gene derived from

XX corynebacterium bacterium, and identifying a homologue of a gene derived from

XX corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino

XX acids, nucleic acids, vitamins, saccharides and organic acids,

XX particularly L-lysine. The present sequence is a protein described in the

XX did not form part of the invention. Note: The sequence data for this patent

XX electronic format directly from the European Patent Office

XX

XX Sequence 661 AA;

XX Query Match 99.0%; Score 3310; DB 4; Length 661;

XX Best Local Similarity 98.9%; Pred. No. 0;

XX Matches 654; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MDHKDLAQRILRDIGGEDNIVAAACATRLRLVLTQTKVDVRSQSLDDPDLKGTETGGM 60

DB 1 MDHKDLAQRILRDIGGEDNIVAAACATRLRLVLTQTKVDVRSQSLDDPDLKGTETGGM 60

QY 61 FOIIVGPGVDVHFPELDDATSKDIASVTEOLKDVANNANMFSRAVKVLADIPVPLPI 120

DB 61 FOIIVGPGVDVHFPELDDATSKDIASVTEOLKDVANNANMFSRAVKVLADIPVPLPI 120

QY 121 LVGGGLMAINNVLVAQDLFGPQSLVEMFPOISGVAEMINIMASAPFAFLPVLVGFTATK 180

DB 121 LVGGGLMAINNVLVAQDLFGPQSLVEMFPOISGVAEMINIMASAPFAFLPVLVGFTATK 180

QY 181 REGNEFLGAGIGMANVPEPTLVNGYDVAATMTAGEMPMSLFGLDVAQGYGTVLPLV 240

DB 181 REGNEFLGAGIGMANVPEPTLVNGYDVAATMTAGEMPMSLFGLDVAQGYGTVLPLV 240

QY 241 VSMILATIEKFLHKLMTADFLITPVLTLLTGFTITAIQPMRWGDLIAHGLQGLY 300

DB 241 VSMILATIEKFLHKLMTADFLITPVLTLLTGFTITAIQPMRWGDLIAHGLQGLY 300

QY 301 DFGPVGGLFGLVYSPVITGTHOSFPPIELFNOGGSFIFATSMANIOGAACLA 360

DB 301 DFGPVGGLFGLVYSPVITGTHOSFPPIELFNOGGSFIFATSMANIOGAACLA 360

QY 361 FFLASEKIKGLAGASGSAVLGITEPAIFGVNRLRMPFYIGITAIIGALLALPDIK 420

DB 361 FFLASEKIKGLAGASGSAVLGITEPAIFGVNRLRMPFYIGITAIIGALLALPDIK 420

QY 421 AVALGAAGFLGVSIDAPDMNMFVCAVVTPTIAGAAIAYGLVYVRNGSIDDPATAP 480

DB 421 AVALGAAGFLGVSIDAPDMNMFVCAVVTPTIAGAAIAYGLVYVRNGSIDDPATAP 480

Db 421 AVALGAGFLGVSIADPMWFLVCAVVTFFIAFGAIAIYGLIYVRNGSIDPDATAAP 480
Qy 481 VPAGTTKAEAEAPAEFSNDSTIIIOAPLTGEAIALSSVSDAMFASGKLGSVAIVPTKGOL 540
Db 481 VPAGTTKAEAEAPAEFSNDSTIIIOAPLTGEAIALSSVSDAMFASGKLGSVAIVPTKGOL 540
Qy 541 VSPVSGKIIVAFPSGIAFAVRTKAEDEGSNDILMHIGFDTVNLNGTHFNPPLKKQGBEYVA 600
Db 541 VSPVSGKIIVAFPSGIAFAVRTKAEDEGSNDILMHIGFDTVNLNGTHFNPPLKKQGBEYVA 600
Qy 601 GELLCEFPDIDAIKAGYEVTTPIVSNYKKTGPVNTYGLGEIEAGANLNLVAKKEAVPAT 660
Db 601 GELLCEFPDIDAIKAGYEVTTPIVSNYKKTGPVNTYGLGEIEAGANLNLVAKKEAVPAT 660
Qy 661 P 661
Db 661 P 661

RESULT 3
AAB6707
ID AAB6707 standard; protein; 468 AA.
XX
AC AAB6707;
XX
DT 09-APR-2001 (first entry)
XX
DE C.glutamincm phosphoenolpyruvate protein #1.
XX
KM Phosphoenolpyruvate; sugar phosphotransferase system; PTS.
XX
OS Corynebacterium glutamicum.
XX
XX WO200102583-A2.
XX
PD 11-JAN-2001.
XX
PF 27-JUN-2000; 2000MO-IB000973.
XX
PR 01-JUL-1999; 99US-0142691P.
XX
PR 23-AUG-1999; 99US-0150310P.
XX
PR 03-SEP-1999; 99DE-01042095.
XX
PR 03-SEP-1999; 99DE-01042097.
XX
PA (BAD1) BASF AG.
XX
PI Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
XX
DR MPI; 2001-080989/09.
XX
XX Corynebacteiium glutamicum nucleic acids encoding phosphoenolpyruvate;
XX
PT sugar phosphotransferase system proteins or their portions, useful for
XX
PT typing or identifying C. glutamicum or related bacteria, and as markers
XX
PT for transformation.
XX
PS Claim 4; Page 101-102; 144pp; English.
XX
XX The present invention relates to Corynebacteiium glutamicum
XX
CC phosphoenolpyruvate; sugar phosphotransferase system (PTS) proteins. The
XX
CC PTS nucleic acids and proteins are useful in the identification of
XX
CC microorganisms which can be used to produce fine chemicals, for
XX
CC modulating fine chemical production in C. glutamicum or related bacteria,
XX
CC the typing or identification of C. glutamicum or related bacteria, as
XX
CC reference points for mapping C. glutamicum genome, and as markers for
XX
CC transformation
XX
SQ Sequence 468 AA;
XX
Query Match 69.8%; Score 2332; DB 4; Length 468;
Best Local Similarity 98.5%; Pred. No. 2,1e-222;
Matches 461; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
Qy 194 MANVFTLVNGIVAAITMTATGEMPMWSLFGLDVAQAGYQGTUVPULVWSHILATIEKFLH 253

Db 1 MANVFPSSLVNGIDVAATTAAGEMPMWSLFGLDVAQAGYQGTUVPULVWSHILATIEKFLH 60
Qy 254 KRLMGTADELITPVLTLITLGTFLTPAIGPAMRWGDLAHGLOGLYDFGPGVGLLFG 313
Db 61 KRLMGTADELITPVLTLITLGTFLTPAIGPAMRWGDLAHGLOGLYDFGPGVGLLFG 120
Qy 314 VSPPIYITGLHOSFPPIELBELFNQGSFTFATBSMANIIOGAACLAVFPLASEKLG 373
Db 121 VSPPIYITGLHOSFPPIELBELFNQGSFTFATBSMANIIOGAACLAVFPLASEKLG 180
Qy 374 GASGVSAVIGITEPAIEFGVNLRLRMPFYIGTAAIGALILAFDITKAVAGAGFLGYV 433
Db 181 GASGVSAVIGITEPAIEFGVNLRLRMPFYIGTAAIGALILAFDITKAVAGAGFLGYV 240
Qy 434 SIDAPDMWFLVCAVVTFFIARFAGAAIYAGLYVRNGSIDPDATAAPVPAAGTTKAEAP 493
Db 241 SIDAPDMWFLVCAVVTFFIARFAGAAIYAGLYVRNGSIDPDATAAPVPAAGTTKAEAP 300
Qy 494 AEFSDNSTIIIOAPLTGEAIALSSVSDAMFASGKLGSVAIVPTKGOLVSPVSGKIIVAP 553
Db 301 AEFSDNSTIIIOAPLTGEAIALSSVSDAMFASGKLGSVAIVPTKGOLVSPVSGKIIVAP 360
Qy 554 SGHAFAPVTRKADGSNDILMHIGFDTVNLNGTHFNPPLKKQGBEYVAGELCEFPDIDAIK 613
Db 361 SGHAFAPVTRKADGSNDILMHIGFDTVNLNGTHFNPPLKKQGBEYVAGELCEFPDIDAIK 420
Qy 614 AAGYEVTTPIVSNYKKTGPVNTYGLGEIEAGANLNLVAKKEAVPATP 661
Db 421 AAGYEVTTPIVSNYKKTGPVNTYGLGEIEAGANLNLVAKKEAVPATP 468

RESULT 4
ABG80325
ID ABG80325 standard; protein; 468 AA.
XX
AC ABG80325;
XX
DT 15-NOV-2002 (first entry)
XX
DE C. glutamicum metabolic pathway (MP) protein #5.
XX
XX
XX Metabolic pathway protein; MP; cell metabolism; amino acid; vitamin;
XX
XX cofactor; nucleotide; nucleoside; trehalose; fine chemical production;
XX
XX organic acid; non-proteinogenic amino acid; purine base; carbohydrate;
XX
XX pyrimidine base; lipid; unsaturated fatty acid; diol; polyketide;
XX
XX aromatic compound; food industry; animal feed; cosmetic industry;
XX
XX pharmaceutical industry; enzyme.
XX
OS Corynebacterium glutamicum ATCC 13032.
XX
XX
XX WO200251231-A1.
XX
XX
XX 04-JUL-2002.
XX
XX
XX 22-DEC-2000; 2000MO-EP013143.
XX
XX
XX 22-DEC-2000; 2000MO-EP013143.
XX
XX
XX (BAD1) BASF AG.
XX
XX Pompejus M, Kroeger B, Zelder O, Schroeder H;
XX
XX MPI; 2002-643289/59.
XX
XX N-PSDB; ABG65346.
XX
XX
XX New metabolic pathway genes of Corynebacterium glutamicum for producing
XX
PT fine chemicals, e.g. lipids, (un)saturated fatty acids, vitamins,
XX
PT cofactors or enzymes used in food, feed, cosmetics or pharmaceutical
XX
PT industries.
XX
PS Claim 18; Page 111-113; 176pp; English.
XX

CC The present invention relates to the isolation of *Corynebacterium*
 CC glutamicum metabolic pathway (MP) proteins, and the polynucleotide
 CC sequences encoding them. The MP proteins are enzymes involved in the
 CC metabolism of molecules important for the normal functioning of cells
 CC (e.g. amino acids, vitamins, cofactors, nucleotides and nucleosides, or
 CC trehalose). The polynucleotide sequences encoding the MP proteins are
 CC useful for producing fine chemicals, particularly organic acids, non-
 CC proteinogenic amino acids, purine and pyrimidine bases, nucleosides,
 CC nucleotides, lipids, (unsaturated fatty acids, diols, carbohydrates,
 CC aromatic compounds, vitamins, cofactors, polyketides and enzymes. The
 CC fine chemicals are useful in the food, animal feed, cosmetic or
 CC pharmaceutical industries. ABG80321-ABG80343 represent the C. glutamicum
 CC MP proteins of the invention
 XX
 SQ Sequence 468 AA;

Query Match 69.8%; Score 2332; DB 5; Length 468;
 Best Local Similarity 98.5%; Pred. No. 2.1e-222;

Matches 461; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 194 NAMVPPTLVNGYDVAAATAGMPWMSLFGDLVAQAQGTGTLPVLVSMIATTEKRLH 253
 DB 1 NAMVPPTLVNGYDVAAATAGMPWMSLFGDLVAQAQGTGTLPVLVSMIATTEKRLH 60
 QY 254 KRLMGADPLTPVLTLLTGTFTFATGPARWGDLLAHGLQGLYDGGPVGGLFGL 313
 DB 61 KRLMGADPLTPVLTLLTGTFTFATGPARWGDLLAHGLQGLYDGGPVGGLFGL 120
 QY 314 VYSPVITGLHQSFPPIELFNOGGSFIFATASMANIAOGAACLAFFLAKSEKLGIA 373
 DB 121 VYSPVITGLHQSFPPIELFNOGGSFIFATASMANIAOGAACLAFFLAKSEKLGIA 180
 QY 374 GASGSVAIVGTEPAIFGNIPLRMPFYIGITTAIGALLLPRIKVALGAAGFLGV 433
 DB 161 GASGSVAIVGTEPAIFGNIPLRMPFYIGITTAIGALLLPRIKVALGAAGFLGV 240
 QY 434 SIDADPMFVLCVAVTFVIAFGAIAVGLYVRNRSIDPDATAPAPAGTTKAEAP 493
 DB 241 SIDADPMFVLCVAVTFVIAFGAIAVGLYVRNRSIDPDATAPAPAGTTKAEAP 300
 QY 494 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSAIVPTKQOLVSPVSGKIVAFP 553
 DB 301 AEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSAIVPTKQOLVSPVSGKIVAFP 360
 QY 554 SGHARAVPTKADGNSVDILMHIGFDTVNLNGTHFNPLKKQDEYKAGEILCEPIDAIK 613
 DB 361 SGHARAVPTKADGNSVDILMHIGFDTVNLNGTHFNPLKKQDEYKAGEILCEPIDAIK 420
 QY 614 AAGYEVTTPIVVSNYKKTGPVNTYIGIEIAGANILNVAKKEAVPATP 661
 DB 421 AAGYEVTTPIVVSNYKKTGPVNTYIGIEIAGANILNVAKKEAVPATP 468

RESULT 5
 AAB66708 standard; protein; 362 AA.
 XX
 AC AAB66708;

DT 09-APR-2001 (first entry)

DE C-glutamicum phosphoenolpyruvate protein #2.

XX Phosphoenolpyruvate; sugar phosphotransferase system; PTS.

XX *Corynebacterium glutamicum*.

XX WO200102583-A2.

XX 11-JAN-2001.

XX 27-JUN-2000; 2000WO-1B000973.

XX

PR 01-JUL-1999; 99US-0142691P.
 PR 23-AUG-1999; 99US-0150310P.
 PR 03-SEP-1999; 99DE-01042095.
 PR 03-SEP-1999; 99DE-01042097.
 XX
 PA (BADI) BASF AG.
 PI
 PI Pompejus M, Kroegeer B, Schroeder H, Zelder O, Habenerauer G;
 DR WPI; 2001-080989/09.
 XX
 XX
 PT *Corynebacterium glutamicum* nucleic acids encoding phosphoenolpyruvate:
 PT sugar phosphotransferase system proteins or their portions, useful for
 PT typing or identifying C. glutamicum or related bacteria, and as markers
 PT for transformation.
 XX
 XX
 PS Claim 6; Page 104-106; 144pp; English.

CC The present invention relates to *Corynebacterium glutamicum*
 CC phosphoenolpyruvate; sugar phosphotransferase system (PTS) proteins. The
 CC PTS nucleic acids and proteins are useful in the identification of
 CC microorganisms which can be used to produce fine chemicals, for
 CC modulating fine chemical production in C. glutamicum or related bacteria,
 CC the typing or identification of C. glutamicum or related bacteria, as
 CC reference points for mapping C. glutamicum genome, and as markers for
 CC transformation
 CC
 SQ Sequence 362 AA;

Query Match 54.0%; Score 1806; DB 4; Length 362;
 Best Local Similarity 99.2%; Pred. No. 2.9e-170;

Matches 359; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 300 YDFGGPVGGLFGLVYSPVITGTHQSFPPIELFNOGGSFIFATASMANIAOGACLA 359
 DB 1 YDFGGPVGGLFGLVYSPVITGTHQSFPPIELFNOGGSFIFATASMANIAOGACLA 60
 QY 360 VFPLAKSEKLGIAAGSASVAVLGITTEPAIFGVNLRMPFYIGITTAIGALLLPDI 419
 DB 61 VFPLAKSEKLGIAAGSASVAVLGITTEPAIFGVNLRMPFYIGITTAIGALLLPDI 120
 QY 420 KAVALGAAGFLGVVSDAPDMVFLVCVAVTFVIAFGAIAVGLYVRNRSIDPDATPA 479
 DB 121 KAVALGAAGFLGVVSDAPDMVFLVCVAVTFVIAFGAIAVGLYVRNRSIDPDATPA 180
 QY 480 PVPAGTTAKAEAPAEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSAIVPTKQ 539
 DB 181 PVPAGTTAKAEAPAEFNSDSTIIQAPLTGEAIALSSVSDAMFASGKLGSAIVPTKQ 240
 QY 540 LVPVSGKIVVAFPSGHAFAVTRKADGNSVDILMHIGFDTVNLNGTHFNPLKKQGEYK 599
 DB 241 LVPVSGKIVVAFPSGHAFAVTRKADGNSVDILMHIGFDTVNLNGTHFNPLKKQGEYK 300
 QY 600 AGEILCEFDIDAIKAGYEVTTPIVVSNYKKTGPVNTYIGIEIAGANILNVAKKEAVPA 659
 DB 301 AGEILCEFDIDAIKAGYEVTTPIVVSNYKKTGPVNTYIGIEIAGANILNVAKKEAVPA 360
 QY 660 TP 661
 DB 361 TP 362

RESULT 6

ADK46609 standard; protein; 627 AA.

XX ADK46609;

XX 20-MAY-2004 (first entry)

XX Streptococcus pneumoniae protein, Seq ID No 3124.

XX Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.

XX

```

XX OS Streptococcus pneumoniae.
XX PN US6699703-B1.
XX PD 02-MAR-2004.
XX PF 26-MAY-2000; 2000US-00583110.
XX PR 02-JUL-1997; 97US-0051553P.
XX PR 12-MAY-1998; 98US-0085131P.
XX PR 30-JUN-1998; 98US-00107433.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Housewart CE;
XX DR MPI; 2004-212399/20.
XX DR N-PSDB; ADK43948.
XX PT New nucleic acid molecules and polypeptides useful for diagnosing,
XX PT preventing and treating pathological conditions resulting from bacterial
XX PT infection, e.g. Streptococcus pneumoniae infection, and in drug
XX PT screening.
XX PS Disclosure; SEQ ID NO 3124; 301pp; English.
XX CC The invention relates to isolated Streptococcus pneumoniae nucleic acids
XX CC and polypeptides. The nucleic acids and proteins are useful for
XX CC diagnosing, preventing and treating pathological conditions resulting
XX CC from bacterial infection, such as S. pneumoniae infection. These may also
XX CC be used for drug screening procedures. The present sequence represents a
XX CC Streptococcus pneumoniae polypeptide of the invention. Note: The sequence
XX CC data for this patent did not appear in the printed specification but was
XX CC obtained in electronic format directly from USPTO at
XX CC seqdata.uspto.gov/sequence.html.
XX SQ Sequence 627 AA;

Query Match 28.9%; Score 965; DB 8; Length 627;
Best Local Similarity 33.4%; Pred. No. 3.1e-86;
Matches 223; Conservative 141; Mismatches 245; Indels 58; Gaps 14;

QY 1 MHKDLAQRILNDIGEDNIVAAAHGATRLRLVLTQTKVDKSLDDDDDKLKTFTGCM 60
DB 1 MNQELIAKKVIALGGRENNVSVAHCATRLRVVQDEEKINKEVINLEKRVQGAFFNSQ 60
QY 61 FOIIVGPGVDVDFEKE--LDATSKDIANVTSQLKDVVANNANWFSRAVKVLADIFVPL 117
DB 61 YQIIFGTGVNKKMYDDEVVLGLFTS-----SKDMLAEYAKQAGNMFQRIKRTGDFVFI 115
QY 118 IPIIVGGGLLMAINNVLAQDLFGPSLVEMFQISVAMINILMASAPFAFLPVLVGFT 177
DB 116 IPIYIVATGIFMGVGRGFNLLEM-----PLPDPATYVQIILDTFTIILPGLVWS 165
QY 178 ATKRFEGNEFLGAGIGMAVFFPLVNGVYAATMTGEMPMMSLFGLDVAQAGYGTVP 237
DB 166 TRRFVFGNPAVGVVLMMLVSGSLPNAWVA--OQGEYVAMNFPGF-IPVGLQGSVLP 221
QY 238 VLVVSVILATIEFKRLMGTAFLITPVLTLTLGFLTFIATIGAMVMVGGDLHLHGQ 297
DB 222 APLIIVGAKFEKAVKVVDPVLDLVTLPVTLVNSIGLFTIGFVHFVENVYILATK 281
QY 298 GYDFGPGVGLLFGLVSPVITVTHQGFPPLELFPNGSGFIF-ATASMANIAQGA 356
DB 282 AILSIFPGFLIGVHOLIVSGVHHFNLELVGLADHANFNPAITTAQTAQGA 341
QY 357 CLAVFLAASEKLGAGAGSVAUIGTERPAIFGVNLRMPFYIGITAAIGALLIL 416
DB 342 TVAVGVKTNPKLTAFPAALSAFLGITEPAIFGVNLRMPFYISLAGAIGGGLASI 401
QY 417 FDIKAVLAGAGLGVISIDAPDMV-----MPLVCAVTFEVIAFGAIAVGLVLR 468

```

```

DB 402 -----LGLAGTNGGNTIIPGMVLVNGQLPQYLLMAVASFALGALTYMGY----- 449
QY 469 NSSIDPATAAPVPACTTAAE-APAEFSNDSTIIQAPLTGEALASSVSAMPASGL 527
DB 450 --EDEVDAATAAKQAEVAEKEEVAAPALQNETLV--TPIVADVVALAVNPDPVFSGAM 505
QY 528 GSGVAIVPTKQGLVSPGSKIIVAFPSGHAFFARTKAEDGSNVDILMHIGPDTVNLNGTH 587
DB 506 GQGIIVAKSQQGVVYAPADAESIAFFTHAFGLKTR---NGAEVLIHIGIDIVSNMG 561
QY 588 FNPILKKQDEVAYAGELCEFDIDAIIKAAGYEVTTPIVSN---YKKTGPVNTYGLGEIA 644
DB 562 FEAKVAQGNKVKXAGDVLGTFDSNKIIAAGLDJDTWIVYVNTADVASVAVAT---GSVAK 618
QY 645 GANLNV 651
DB 619 GDAVIEV 625

RESULT 7
ADR94615
ID ADR94615 standard; protein; 643 AA.
AC ADR94615;
DT 16-DEC-2004 (first entry)
DE Novel S. pneumoniae protein sequence, SEQ ID 3250.
DE Meningitis; bacteraemia; pneumonia; otitis media; vaccine;
KM bacterial infection.
XX OS Streptococcus pneumoniae.
XX PN US6800744-B1.
XX PD 05-OCT-2004.
XX PF 30-JUN-1998; 98US-00107433.
XX PR 02-JUL-1997; 97US-0051553P.
XX PR 12-MAY-1998; 98US-0085131P.
XX PA (GENO-) GENOME THERAPEUTICS CORP.
XX PI Doucette-Stamm LA, Bush D;
XX DR MPI; 2004-697205/68.
XX DR N-PSDB; ADR92012.
XX PT New isolated nucleic acid encoding a Streptococcus pneumoniae
XX PT polypeptide, useful for diagnosing, preventing and/or treating
XX PT pathological conditions resulting from the bacterial infection.
XX PS Disclosure; SEQ ID NO 3250; 151pp; English.
XX CC The invention relates to an isolated nucleic acid comprising a sequence
XX CC encoding a Streptococcus pneumoniae ADR91366polypeptide, or its
XX CC fragments, with any of 9 fully defined sequences (appearing as ADR94308,
XX CC ADR94489, ADR94800, ADR94837, ADR94969, ADR95253, ADR95642, ADR91705,
XX CC ADR96079) or any of the fully defined sequences appearing as ADR91705,
XX CC ADR91886, ADR92197, ADR92234, ADR93039, ADR93079, ADR92366, ADR92650 or
XX CC ADR93476 or at least 20 or 30 consecutive nucleotides of the nucleotide
XX CC sequences, or at least 40, 60 or 300 consecutive nucleotides, which is
XX CC hybridizable under high stringency conditions to the nucleotide sequence.
XX CC The nucleic acids and proteins are chosen from 5206 disclosed sequences.
XX CC Also included are a recombinant expression vector comprising the isolated
XX CC nucleic acid cited above operably linked to a transcription regulatory
XX CC element, a cell comprising the recombinant expression vector and a probe
XX CC comprising at least 20 consecutive nucleotides of the nucleotide
XX CC sequences as cited above. The methods and compositions of the present
XX CC invention are useful for the diagnosis, prevention and/or treatment of
XX CC pathological conditions resulting from bacterial infection by

```


Query Match	28.6%;	Score 957;	DB 5;	Length 617;
Best Local Similarity	35.1%;	Pred. NO. 1.9e-85;		

SQ Sequence 627 AA;

Query Match	28.6%;	Score 956;	DB 4;	Length 627;
Best Local Similarity	33.3%;	Pred. No. 2.5e-85;		
Matches 222;	Conservative 141;	Mismatches 246;	Indels 58;	Gaps 14

```

Oy 1 MDHOKLABIENDIGGEDENI VAAHOCATRLVLT.KOTKOVDRGSLDDDDPLKOTFFEGM 60
Db 1 MNDGI AKKVALDAGRENNVSVAHATILRLVWVKOBEKINKENIENLEKVGQAFRNSGQ 60
Oy 61 FOIIVPGDVBHVFKE---LDDATSKDIAVSTEQLDXDVANNANWFSRAVKVLADIFVPL 117
Db 61 YQIIFGTGVNKMVDEVVVLGLPTS-----SKDKMAEYBAKQGNWFQRAIRFTGDFVPL 115
Oy 118 IPILVGGGLMAINNVLVAODLFGPQSLVEMFPQISGVAEMINLMASAPPAFLPVLVGT 177
Db 116 IPIVATGTFEMVGRGJFNALEM-----PLPGDFATYQIOLDITDAFIIIFGLVWS 165
Oy 178 ATKRFGEFELGAGIGMAMVFFTLVNGVUAATMTAGEBPMMSLFGJLDVAQAQGYCTVLP 237
Db 166 TFRVFGNPAVGI VLGMMLVSGSLPRAMAVA---QGEVLTANNFFBF-IPVGLQSSVLP 221
Oy 238 VLVVSMILATIEKFHLKRLMGTADELITVLTLLTGFTFIALGPMERWGDLLAHGLO 237
Db 222 AFIIIGVAKKEKAVRKVPDVIDLVTPEPVLVWSIIGLFVIGVFPHVENYIILATK 281
Oy 298 GLYDEGCPVGGELFGLVYSPVITGTGHQSFPIETLFPNGGSEFIE-ATASMNIHQGA 356
Db 282 AILSPFELGGLTIGVQVOLIYSGVHNI.FNILEVOLLADHANPNAIITTAAMTQGA 341
Oy 357 CLAVEFLAKSEKULGAGVSAVLGITPEAIFGVNLRLRMFXYIGICTAIGALIAL 416
Db 342 TVAVGVTKNPKTLTILAPALSAFAGITEPAIFGVNLRFKRFELSLAIGIGGLASI 401
Oy 417 FDIKVALGAGFLGVNSIDAPDMV-----MPLVCAVYTFVIAFGAIALYGLYVR 468
Db 402 -----JGLMGTNGITITPGMTLVNGGOL.PQYLLMAVSPALGFALTVMGY----- 449
Oy 469 NSGIDPDATPAVPVPAATTAAEAE-APAEFSNDSTIIQADLTGEALISSVSDAMFASGL 527
Db 450 --EDEVDATAAKQAEVAEKEKEVAPALQNEILV--TLIVGVVALADVNDVFSSGM 505
Oy 528 GSGVAIVTPKQOLVSPVSGKIVVAPPSGIAFAVRTAEDGSNVDILMHIGFTVINLNGH 587
Db 506 GOGIAVKSOGVVVALDAEVAISIAFPTGA.FGLKTR---NGAEVLI.IHGIDIVSNMGOG 561
Oy 588 FNP.LKQDDEVVAGELCEFDIDA.KAAGYEVTTPVNSN---YKKTGPNTGLGEIEA 644
Db 562 FEAKYAKQGNKYAKAGVGLGTFDSNKLTAAGLDDITMIVITNTADYASVAEYAT---GSVK 618
Oy 645 GANLINV 651
Db 619 GDAVIEV 625

```

RESULT 12
ABU02378
ID ABU02378 standard; protein; 655 AA
XX
XX ABU02378;
XX

DT	23-OCT-2003	(revised)
DT	11-FEB-2003	(first entry)
XX		
DE		S. pneumoniae type 4 strain protein from coding region #1956.
XX		
KW		Bacterial meningitis; pneumonia; sepsis; otitis media; ear infection;
KW		antiflammatory; antibacterial; immunostimulant; auditory; respiratory;
KW		gene therapy; vaccine.
XX		
OS		Streptococcus pneumoniae; type 4 strain.
XX		
PN	MO200277021-A2.	
XX		
PD	03-OCT-2002.	
XX		
FE	27-MAR-2002; 2002NO-IB002163.	
XX		
PR	27-MAR-2001; 2001GB-00007658.	
XX		
PA	(CHIR-) CHIRON SPA.	
PA	(GENO-) INST GENOMIC RES.	
XX		
PI	Masignani V, Tettelin H, Fraser C;	
XX		
DR	WPI; 2003-040579/03.	
DR	N-PSDB; ABX07668.	
XX		
PT		New proteins and nucleic acid molecules from Streptococcus pneumoniae,
PT		useful as medicaments for treating or preventing a disease or infection
PT		due to Streptococcus bacteria, such as pneumonia, sepsis, otitis media or
PT		ear infection.
XX		
PS	Claim 1; SEQ ID NO 3912; 56pp; English.	
XX		
CC		The invention relates to a protein comprising or having at least 50t
CC		identity to any of the 2469 amino acid sequences, identified in the
CC		specification (available on a computer readable format), or its fragment,
CC		expressed from 2469 of 2489 identified DNA coding regions from the
CC		Streptococcus pneumoniae type 4 strain genomic sequence appearing as
CC		ABSS6454. Also included are an antibody which binds one of the proteins,
CC		treating a patient by administering the protein, DNA or antibody (in a
CC		composition), a kit comprising first and second primers, which are the
CC		nucleic acid cited above or fragments between nucleotides 8-100 of a
CC		sequence not defined in the specification, for amplifying a target
CC		sequence contained within a Streptococcus nucleic acid sequence, where
CC		and the first primer is substantially complementary to the target sequence
CC		and the second primer is substantially complementary to the target sequence
CC		the target sequence, and where the parts of the primers having
CC		substantial complementarity define the termini of the target sequence to
CC		be amplified, assay comprising contacting a test compound with the
CC		protein, and determining whether the test compound binds to the protein
CC		and a Streptococcus pneumoniae bacterium, where one or more genes
CC		encoding the proteins has been rendered inactive. The proteins, nucleic
CC		acid molecules, antibody and compositions are useful as medicaments for
CC		treating or preventing a disease or infection due to streptococcus
CC		bacteria, particularly S. pneumoniae, such as pneumonia, sepsis, otitis
CC		media or ear infection. They are also useful in developing vaccines,
CC		diagnostics and antibiotics. The methods are useful for identifying
CC		immunodominant proteins. The present sequence is one of the 2469 proteins
CC		expressed by the identified coding regions from the genomic sequence.
CC		Note: The sequence data for this patent did not form part of the printed
CC		specification, but was obtained in electronic format directly from WIPO
CC		at ftp.wipo.int/pub/published_pct_sequences. (updated on 23-OCT-2003 to
CC		standardise int field)
XX		
XX	Sequence 655 AA;	

```

Query Match      28.4%; Score 949; DB 6; Length 655;
Best Local Similarity 34.9%; Pred. No. 1.3e-84;
Matches 234; Conservative 128; Mismatches 253; Indels 56; Gaps 19

7 AGRIRLDIGGNNIYAAAHGCAIRLRRLVTKDTKVDKDSLDDDDDKAGTETGMPQITNG 66
| : : : | : : : : | : : : : | : : : : | : : : : | : : : : |

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 21:44:36 ; Search time 43 Seconds
(without alignments)
1147.512 Million cell updates/sec

Title: US-10-019-284b-2

Perfect score: 3342

Sequence: 1 MDHKLADRLDIDGEDNI.....IEAGNLANVAKKEAVPATP 661

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*

2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*

3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*

4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*

5: /cgn2_6/prodata/1/1aa/6C_COMB.pep:*

6: /cgn2_6/prodata/1/1aa/6D_COMB.pep:*

7: /cgn2_6/prodata/1/1aa/6E_COMB.pep:*

8: /cgn2_6/prodata/1/1aa/6F_COMB.pep:*

9: /cgn2_6/prodata/1/1aa/6G_COMB.pep:*

10: /cgn2_6/prodata/1/1aa/6H_COMB.pep:*

11: /cgn2_6/prodata/1/1aa/6I_COMB.pep:*

12: /cgn2_6/prodata/1/1aa/6J_COMB.pep:*

13: /cgn2_6/prodata/1/1aa/6K_COMB.pep:*

14: /cgn2_6/prodata/1/1aa/6L_COMB.pep:*

15: /cgn2_6/prodata/1/1aa/6M_COMB.pep:*

16: /cgn2_6/prodata/1/1aa/6N_COMB.pep:*

17: /cgn2_6/prodata/1/1aa/6O_COMB.pep:*

18: /cgn2_6/prodata/1/1aa/6P_COMB.pep:*

19: /cgn2_6/prodata/1/1aa/6Q_COMB.pep:*

20: /cgn2_6/prodata/1/1aa/6R_COMB.pep:*

21: /cgn2_6/prodata/1/1aa/6S_COMB.pep:*

22: /cgn2_6/prodata/1/1aa/6T_COMB.pep:*

23: /cgn2_6/prodata/1/1aa/6U_COMB.pep:*

24: /cgn2_6/prodata/1/1aa/6V_COMB.pep:*

25: /cgn2_6/prodata/1/1aa/6W_COMB.pep:*

26: /cgn2_6/prodata/1/1aa/6X_COMB.pep:*

27: /cgn2_6/prodata/1/1aa/6Y_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	965	28.9	627	4	US-09-583-110-3124 Sequence 3124, Ap
2	965	28.9	643	4	US-09-107-433-3250 Sequence 3250, Ap
3	960.5	28.7	496	4	US-09-134-000C-5822 Sequence 5822, Ap
4	935	28.0	655	4	US-09-583-110-2958 Sequence 2958, Ap
5	923	27.6	628	4	US-09-107-532A-5288 Sequence 5288, Ap
6	875.5	26.2	656	4	US-09-489-039A-8212 Sequence 8212, Ap
7	855	25.6	670	4	US-09-134-000C-4606 Sequence 4606, Ap
8	826	24.7	631	4	US-09-107-532A-3902 Sequence 3902, Ap
9	812	24.3	483	4	US-09-489-039A-12018 Sequence 12018, A
10	802.5	24.0	634	4	US-09-543-681A-7019 Sequence 7019, Ap
11	796.5	23.8	612	4	US-09-583-110-2911 Sequence 2911, Ap
12	773	22.1	583	4	US-09-107-532A-6611 Sequence 6611, Ap
13	678	20.3	530	4	US-09-107-433-5169 Sequence 5169, Ap
14	650.5	19.5	481	4	US-09-543-681A-5156 Sequence 5156, Ap
15	644	19.2	590	4	US-09-107-532A-6004 Sequence 6004, Ap
16	642.5	19.2	475	4	US-09-489-039A-8977 Sequence 8977, Ap
17	578.5	17.3	435	4	US-09-107-433-5184 Sequence 5184, Ap
18	558	16.7	526	4	US-09-134-000C-4715 Sequence 4715, Ap
19	537	16.1	316	4	US-09-134-000C-5482 Sequence 5482, Ap
20	528.5	15.8	243	3	US-09-134-001C-4114 Sequence 4114, Ap
21	511	15.3	478	3	US-09-134-001C-5065 Sequence 5065, Ap
22	510	15.3	460	4	US-09-543-681A-5773 Sequence 5773, Ap
23	495.5	14.8	427	4	US-09-107-532A-5530 Sequence 5530, Ap
24	462.5	13.8	482	4	US-09-489-039A-9909 Sequence 9909, Ap
25	437.5	13.1	482	4	US-09-107-532A-6644 Sequence 6644, Ap
26	417.5	12.5	479	4	US-09-489-039A-10571 Sequence 10571, A
27	401	12.0	527	4	US-09-634-238-368 Sequence 368, App

28	329	9.8	696	3	US-09-134-001C-4296 Sequence 4296, Ap
29	315	9.4	656	4	US-09-107-532A-5083 Sequence 5083, Ap
30	312	9.3	241	4	US-09-134-000C-4040 Sequence 4040, Ap
31	310.5	9.3	726	4	US-09-583-110-2782 Sequence 2782, Ap
32	310.5	9.3	726	4	US-09-107-433-3906 Sequence 3906, Ap
33	306	9.2	635	4	US-09-489-039A-13395 Sequence 13395, A
34	298	8.9	688	4	US-09-543-681A-5236 Sequence 5236, Ap
35	296.5	8.9	585	4	US-09-107-532A-4366 Sequence 4366, Ap
36	271.5	8.1	470	4	US-09-107-532A-5798 Sequence 5798, Ap
37	267	8.0	170	3	US-09-134-001C-5134 Sequence 5134, Ap
38	256	7.7	196	4	US-09-543-681A-7659 Sequence 7659, Ap
39	249	7.5	315	4	US-09-134-000C-3657 Sequence 3657, Ap
40	244	7.3	172	4	US-09-489-039A-9162 Sequence 9162, Ap
41	230	6.9	686	4	US-09-252-991A-18115 Sequence 18115, A
42	224	6.7	197	4	US-09-107-433-4682 Sequence 4682, Ap
43	197.5	5.9	463	4	US-09-134-000C-3598 Sequence 3598, Ap
44	183.5	5.5	98	3	US-08-936-165A-419 Sequence 419, App
45	181.5	5.4	153	4	US-09-107-532A-4365 Sequence 4365, Ap

ALIGNMENTS

RESULT 1	US-09-583-110-3124	Sequence 3124, Application US/09583110
Patent No. 6699703		
GENERAL INFORMATION:		
APPLICANT: Lynn Doucette-Stamm et al.		
TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus		
TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics		
FILE REFERENCE: PAT00-07A		
CURRENT APPLICATION NUMBER: US/09/583,110		
PRIOR APPLICATION NUMBER: 2000-05-26		
PRIOR FILING DATE: 1998-06-30		
PRIOR APPLICATION NUMBER: US 60/085,131		
PRIOR FILING DATE: 1998-05-12		
PRIOR APPLICATION NUMBER: US 60/051,553		
PRIOR FILING DATE: 1997-07-02		
NUMBER OF SEQ ID NOS: 5322		
SEQ ID NO 3124		
LENGTH: 627		
TYPE: PRT		
ORGANISM: Streptococcus pneumoniae		
US-09-583-110-3124		
Query Match	28.9%	Score 965; DB 4; Length 627;
Best Local Similarity	33.4%	Pred. No. 2.9e-90;
Matches 223; Conservative 141; Mismatches 245; Indels 58; Gaps 14;		
1 MDHKLADRLDIDGEDNIYAAHCATRLVLTDTXDVDRQSLDDPDLKGTETGCM 60		
1 MNNQSLAKVLDALGREGVNSVAHCATRLRMVDEKINEVEENLEKVGAFNSQ 60		
61 FOIIVGPDVDFVKE---LDDATSKDIAVSTEQIKDVVANNANFSAVKTLADIFPPL 117		
61 YOIIFGTGVNKMVDEVVVLGLPTS-----SKDDKAEVAKQGNMFORAIRFGGVFVPI 115		
118 IPIIVGGGLIIMANNVLAQDLPQSPVEMRQISGVAEMINIMASAPFAPLPVLYGT 177		
116 IPIVATGELFMVGRGIFNLEH-----PLPGRFATYQIITDRAFIILPGLVWS 165		
178 ATKRGNEFLGAGIGMAVPTLVNGDVATMTAGEMPWMSLFGLDVAQGYGQTVLP 237		
166 TRVRGNGNPAVGIIVGMVLVSGSLPNAMVA---GGGETLANPFGF-IPVVGLOGSVLP 221		
238 VLVSMILATIEKFLAKLMGTADLPVLTLLTGLTFTFAIGPANKWGDLLAHGQ 297		
222 AFIIVGAKKFEKAKRVKVPDVIDLVTPFVTLVMSITGLFVIGPVHVNNYILATK 281		
298 GLYDGGPVGGLIFGLVNSPIYITGLHOSFPFIELEFNQGSFF-ATASMANIAGAA 356		

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5607
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5288:
SEQUENCE CHARACTERISTICS:
LENGTH: 628 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (8) LOCATION 1...628
SEQUENCE DESCRIPTION: SEQ ID NO: 5288:
US-09-107-532A-5288

Query Match 27.6%; Score 923; DB 4; Length 628;
Best Local Similarity 32.3%; Pred. No. 6,2e-86;
Matches 215; Conservative 134; Mismatches 253; Indels 64; Gaps 12;
QY 1 MDKDLAQRILRDIGEDNIVAAGCATRLVLKDTKDVRSQSLDDPDLKTEETGGM 60
DB 8 MDMQAVGRVWEAVAGQKVKSLVCATRLRFLKDESLADTKEDPDVIGVWESGGQ 67
QY 61 FOIIVPGDVHVFELDDATSKDIASVTEQLKDVVANNANNSRAVVKLADIFVPLPI 120
DB 68 YQVAVIG-SNVADVYOSIYD--EEGLAVS---EDGKESKNILNRIIDISSIFPFGLA 120
QY 121 LVGGGLMANNVVAQDLFGPSLVEMFPQISGVAEMINLMAAPFAFLPVLVGFTATK 180
DB 121 MAAAGVLKGFSL-----ATVAGMILTADSGTYQIIFAAADGVFTFLPMLAFTAK 171
QY 181 REGNEFLGAGIGMANVFPFLVNGYDVAATMTAGEMPWMSLFGLDV--AQAGYQGVLPV 238
DB 172 KFKAQFLAVAIAMLVLP-----ATTAVAGAGEAISFGIPIVILSPSGYTSVPI 223
QY 239 LVSWILATIEKFLHKLMTADFLITPVLLTLLGFLTFIAGPAMRWGDLAHGLOG 298
DB 224 ILAVWQSKLEPFVKVIVPOLQMLIVPLVAVVWVPLTFELAGPITVAGNALGLEN 283
QY 299 LVDFGPGVGLFGLVYSPVITGLHQSFPPIE-LELFNQGSGSFIFATAMANIQAAC 357
DB 284 IYGFSEYIAGLIMGSMQVFMFGMHWGVPFIMFLITEQYGFVILPMLIPAILAOGGA 343
QY 358 LAVFLAKSEKLGKLAGAGVSAVLGITPEPAIFGVNLRMPFYIGITGAALGALIALF 417
DB 344 LAVALRTKQTKLRALGISSTVSLRGTETPYGVTLPLKPFIAACISGGIGGAILIGS 403
QY 418 DIKAVA-----LGAAGFLGVNSIDAPDMVFLVCAVVTFTVIAFGAALYGLIVRRNSI 471
DB 404 GYKAFSSSLVSLITPTFTINTVDGVSNTVAIVIAIGFVIAFVLTLLIGFDEQODQ 463
QY 472 IDPDATTAAPVPAAGTTAAEAPAEFENDSTIIQAPLTGAIALSSVSDMPFSGKLGSGV 531
DB 464 EE-----NKHANAGBPITTSABHTLKSPLTGKVLPISEVPDGVFSSGVWGKGI 510
QY 533 AIVPTGQLVSPVSGKIVVAFPSGAFAVRTKAEDSNDIIMHIGFDVNLNGTHFNPL 591
DB 511 AIDPEVGEIIVAPADGKITITTFPGHAGVGLTT--TDGA--EILIHIGMDVVELNGNGFELL 566

QY 592 KKQGEVAGELTCEPDIDAIIKAAGYEVTYIVSNYKKGTPVNTYGLGEIEAGANILNV 651
DB 567 VKQGDVYAGDGLIRFDIEAIRAAGYSVITVITN-----TDAPRDIETL 612
QY 652 AKKEAV 657
DB 613 DQKEII 618
RESULT 6
US-09-489-039A-8212
Sequence 8212, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709,2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 8212
LENGTH: 656
TYPE: PR
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8212

Query Match 26.2%; Score 875.5; DB 4; Length 656;
Best Local Similarity 32.7%; Pred. No. 5.4e-81;
Matches 216; Conservative 124; Mismatches 271; Indels 49; Gaps 13;
QY 1 MDKDLAQRILRDIGEDNIVAAGCATRLVLKDTKDVRSQSLDDPDLKTEETGGM 60
DB 37 MEYKALADILNRVGGKENIVSLVHCATRLFKLKDNGKADAEGLKANPVMVESGGQ 96
QY 61 FOIIVPGDVHVFELDDATSKDIASV--TEQLKDVVANNANNSRAVVKLADIFVPLI 118
DB 97 FOVAVIG--NHV-HDVLAARQEXGSDSEPVAEERKAGSVLSQDLIDISGLFTFPI 151
QY 119 PILVGGGLMANNVVAQDLFGPSLVEMFPQISGVAEMINLMAAPFAFLPVLVGFTA 178
DB 132 GVMATGLIKKLIALAVTCMLTPE-----QTYKIMFASDAPFFFPFLFGTGA 202
QY 179 TKREGNEFLGAGIGMANVFPFLVNGYDVAATMTAGEMPWMSLFGLDVVAQAGYQGVLPV 238
DB 203 GKKEGNEFFISWVIGALTLHPMTQAFSAQAPGA--VEHFLGI PVTFINYSVPI 259
QY 239 LVSWILATIEKFLHKLMTADFLITPVLLTLLGFLTFIAGPAMRWGDLAHGLOG 298
DB 260 ILASVWSCMLERKSNALLPSSMKOFFSPALCLAVVPLTFVIGVATWLSHLANGYOF 319
QY 299 LVDFGPGVGLFGLVYSPVITGLHQSFPPIE-LELFNQGSGSFIFATAMANIQAAC 357
DB 320 IYAFPMIAGAVLGMQVCIIFGLHGLVPLMINMNVVLGHDSMLPIILRAVIRQVAV 379
QY 358 LAVFLAKSEKLGKLAGAGVSAVLGITPEPAIFGVNLRMPFYIGITGAALGALIALF 417
DB 380 LGIFLATDARQVLAGSAFSAFLGITPEPAIYGLTLPLRBPFIGVAGAIGAITAFS 439
QY 418 DIKAVAGAGFLGVNSIDAP---DMVNF--LVCAVVTFTVIAFGAALYGLIVRRNSI 472
DB 440 NSYAFSFLPNIFFPDAQMIPEGIDASVWGGLIGTGVAFVLA CVLTFPAGL----- 490
QY 472 DPDATTAAPVPAAGTTAAEAPAEFENDSTIIQAPLTGAIALSSVSDMPFSGKLGSGV 532
DB 491 PRGSAP-----GAVTVAPAS-AND--ILAPSSGVILBOVPDSTFASGLIGKVA 539
QY 533 IVPITGQLVSPVSGKIVVAFPSGAFAVRTKAEDSNDIIMHIGFDVNLNGTHFNPL 592
DB 540 IIPAVGQVIAAPPGVAVSLFQTKAIGL-----QDSGIELIHVIGIDIVKLDGVFPFAHV 595

Query Match 22.1%; Score 737; DB 4; length 583;
Best Local Similarity 30.6%; Pred. No. 8,98-67;
Matches 180; Conservative 119; Mismatches 239; Indels 50; Gaps 13;

Cy 87 VSTEQLDVVANNMPSRAVKYLAADIFVPLIPITLVGGGLMAINYLYVA---ODLPGQ 143
 :
Db 7 VSTEOGKAARKONHAPORAIAYVALEIFETPLIPAIIYVGCHILGRNVLTEGIQPSLEG-- 64

QY 144 SLVEMPQISGVAEMINIMASAPFAFLPVLTGFTATKREGNEFLGAGIMAVPPTLVN 203
 DB 65 TIVESHKFMNWNVAIFMLPGEBALFHPLEPGITMSIAKKGTTOILSTIVIGITLVSPOLN 124
 QY 204 GYDVAATMTAGEMPMWSLFGLDVAQYOGTVLPVLVSMILATEKFLHKRLMTADFL 263
 DB 125 AVSVAAT-AAADIPFDFGPAQVOMGYQAVIPALALAGFMALYEIFRKXIPQISIM 183
 QY 264 ITPEVTLTLTGFTLFAIGP---AMRWGDLIAGLQGLYDFGGVGLLGLVYSPV 319
 DB 184 FVPLFSLPVLAAHYLTSPIGWTGSMISINVTGTLTSSIXW---LFSAVGFLYAPLV 240
 QY 320 ITGLHOSPEPIELEFNQGSF-IFATASMANIAGAACLAFLAK-SEKXKGLAGAG 377
 DB 241 ITGLHMMNNAIDMQLIADGSGSTLMPMILSNIAQSALVALIIFLHKGKKEQISIPM 300
 QY 378 VSAVIGITEPAIFGVNLRMPFYIGITPAIGALIALFDIKAVLAGAAGLVGVSIDA 437
 DB 301 ISCYIGTEPAMGINKLVYPPVAMVSGLAGFANLMGVRAANIGVGLPGILAIQA 360
 QY 438 PDMWFLVCAVTVTVIAPGALAYGLYVRNG---SIDP-DATAPVPAGTTAKAABA 492
 DB 361 ETWVPEIIMIAVILPFGILTIIF---RQGIINKIDPAVPENADVOLQATANGATAT 415
 QY 493 PAEFSNDSTIIQAPLTGEAIA-----LSSVSDAMPASGLSGSVAILPTK 537
 DB 416 POSFEVSA-----TGNVATKKEITFAVAAGTITKEITVNDVFSQKMGOGYAVEPSN 469
 QY 538 GOLVSPVSGKIVAFPSGHAFAVRTKADGSDNVDIMHIGFDIVNLNTHFNPLKQGD 597
 DB 470 GKYYAVNKNVTSVFETKHAIGILS---NEGLEVIVHMGDLTVELKGVFVNVFVGEVL 525
 QY 598 VYAGELCEPDIDAIKAAGYEVTTPVIVSNYKKTGVNTYTGCEIAG 645
 DB 526 VPEETLIAEMDPEIEQAGKTDIIVALTNNKEVGLSIDQGLVPRG 573
 RESULT 13
 US-09-107-433-5169
 / Sequence 5169, Application US/09107433
 / Patent No. 6800744
 / GENERAL INFORMATION:
 / APPLICANT: Lynn A Doucette-Stamm and David Bueh
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
 / SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
 / THERAPEUTICS
 / NUMBER OF SEQUENCES: 5206
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 / STREET: 100 Beaver Street
 / CITY: Waltham
 / STATE: Massachusetts
 / COUNTRY: USA
 / ZIP: 02154
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: CD-ROM ISO9660
 / SOFTWARE: <Unknown>
 / OPERATING SYSTEM: <Unknown>
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/09/107,433
 / FILING DATE: 30-Jun-1998
 / PRIOR APPLICATION DATA:
 / APPLICATION NUMBER: 60/ 085131
 / FILING DATE: May 12, 1998
 / APPLICATION NUMBER: 60/051553
 / FILING DATE: July 2, 1997
 / ATTORNEY/AGENT INFORMATION:
 / NAME: Ariniello, Pamela Deneke
 / REGISTRATION NUMBER: 40,489
 / REFERENCE/DOCKET NUMBER: GTC-011
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: (781) 893-5007

; TELEFAX: (781) 893-8277
 ; INFORMATION FOR SEQ ID NO: 5169:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 530 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHEICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Streptococcus pneumoniae
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (B) LOCATION 1...530
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5169:
 US-09-107-433-5169
 Query Match 20.3%; Score 678; DB 4; Length 530;
 Best Local Similarity 33.7%; Pred. No. 9.2e-61;
 Matches 168; Conservative 101; Mismatches 184; Indels 46; Gaps 15;
 QY 174 VGFATKFGGNEFGAGIGAMVPTLVNGYDVAATMTAGEM--MMSLGLDVAQAGY 231
 DB 50 ITWSVRKQGTNOIGIVIGICVSPOLLNVAVASTPAADIAANWVNFQFTVNRIGY 109
 QY 232 OGTVLPVLVSMILATIEKFLHKRLMTADFLITPEVTLTLTGFTLFAIGPAMRWGDL 291
 DB 110 QAGVTPALLAGLSYLEIFRKXIPQISIMFVPLSLIALLIATHTVLDP---IGWT 165
 QY 232 LAHQGLQGLYDFG--GPVG---GLLEGLVYSPVITGTHQSPPEIELEF--NOGGSFIPAT 345
 DB 166 IGGGSLSVLAGLTGPVAKMLFGALFYAPVITGTHHMTNAIDTQLIADAGGTALWPM 225
 QY 346 ASMANIAGAACLAFLPAKSEKLG--LAGSGVSAUVGITPEAFGNILMMPFYIGI 404
 DB 226 IALSNIAQGSAAVFAFYFMRHDECAQVSLPATTISAVIGTEPFLFGVNVKXIYFVAGM 285
 QY 405 GTAAGCALIALFDIKAVLAGAAGLVGVSIDAPDMWF---IYCAVTVTVIAPGALIA 460
 DB 286 TGSALAGMLSTFNTAASIGIGLPGILISTOPQYMLFPAGTMLVAYVPMILTT----- 340
 QY 461 YGLYVLRNG---SIDPDATAAPVAGTTKAE--AEAPAEFSNDSTIIQAPLTGEAIALS 515
 DB 341 ----PFRKAGLPETKIEGDTNLQAEFVQAEAEFVSHBEVELT--SVEIISPLTGQVKELS 394
 QY 516 SVSDAMPASGLSGSVAILPTKQGLVSPVSGKIVAFPSGHAFAVRTKADGSDNVDIMH 575
 DB 395 QATDVPFASGVWGGLVIEPSQGLTSPVNGTVTLFPTKHAIGI--VSDG--VELLIH 450
 QY 576 IGFDIVNLNTHFNPLKQGDVKGAGELCEPDIDAIKAAGYEVTTPVIVSNYK----- 630
 DB 451 IGMDIVGLDGKGFESLVQGDHVIYVGOQLIRFDMDIVTAAGLVETPVIITNQDAYTATI 510
 QY 631 TGPVNTYGLGEIEAGANLL 649
 DB 511 TGTPTPT-----TIQAGASIM 525
 RESULT 14
 US-09-543-681A-5156
 / Sequence 5156, Application US/09543681A
 / Patent No. 6605709
 / GENERAL INFORMATION:
 / APPLICANT: GARY BRETON
 / TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
 / FILE REFERENCE: 2709.1002-001
 / CURRENT APPLICATION NUMBER: US/09/543,681A
 / PRIOR FILING DATE: 2000-04-05
 / PRIOR APPLICATION NUMBER: US 60/128,706
 / PRIOR FILING DATE: 1999-04-09
 / NUMBER OF SEQ ID NOS: 8344
 / SEQ ID NO 5156
 / LENGTH: 481


```

; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5156

Query Match      19.5%; Score 650.5; DB 4; Length 481;
Best Local Similarity 31.7%; Pred. No. 5.4e-58;
Matches 150; Conservative 111; Mismatches 197; Indels 15; Gaps 10;

QY      8 QRLRDIGEDNIVAAAHGATRLRLVLTQTKVDVROSDDDDPKGTFTGGMFOIIVG 67
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      19 EELITLVGSDNIVSYSHCLTRLRFLVLPNLADIDIKIKISFVKCCFNNAQGFQYIVGM 78
QY      68 GVDVHFKELDQATSDIAVSTEQLDVYVANNANWFSRAVKYADI FVPLPILVGGGL 127
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      79 -DVDIYKLLIDKLSLK-AKSKODISDAKONGSQKLANIAELFVPLPILVGGGL 136
QY      128 MAINVLVAODLFGPQSLVEMFPQISGVAEMINLMASAPAFPLVVGFTATRFQGNFR 187
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      137 LGLRNVIGEMPIISDNKPLTFYFVPLTPIDFLMPCBEAIFHFLPVSICMSSVKKMGSTRA 196
QY      188 LGAGIGMANVFPVLVNGYDVAAITMTAGEMP-MMSLGL-DVAGAGTGYVLPVLVSWIL 245
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      197 LGLVIGITIVSPQIMANVNGS-----QLPEVWN-FGLFTIEKVGYOQVIPSILGLFL 250
QY      246 ATIEKFLHKLMTADFLITPVLTLLTGFLTFIAGPMRWGDLAHLGLGLY--DYG 303
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      251 GMEFTKRRFIPYLVKLVIVPIVTLVAVFAHVILGPIGRSIGDGLAQLIKPLMGDF- 309
QY      304 GPVGLLFGLVSPVITVGHOSFPIEELF-NOGGSFIFATASNANIQAACLAVER 362
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      310 APIGSTIFAFVSPWITGHTHTLAIIDQMOSYGSTPIWPIALISNIAQASAVVGIII 369
QY      363 LAKSEKLGAGAGSVSAVLGTEPAIFGVNLRMPFIYIGTAAIGALLAFEDIKAV 422
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      370 VSSKLNQEOETITPAIASAYLGIVTEPAMYGNLKYGFPMLCAMIGALLGICLNVLN 429
QY      423 ALGAGFLGVSTIDAPDMWFLVCAVTFPIARGAALVGLVVRNGSIDPD 475
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      430 GIGVGLPGILSIQPTYWLVLTAMVAVIPILGLITLVYRY-KERAGTLQVD 481

RESULT 15
US-09-107-532A-6004
; Sequence 6004, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneka
; REGISTRATION NUMBER: 40,489

```

```

; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 6004:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 590 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (b) LOCATION 1...590
; SEQUENCE DESCRIPTION: SEQ ID NO: 6004:
US-09-107-532A-6004

Query Match      19.3%; Score 644; DB 4; Length 590;
Best Local Similarity 26.6%; Pred. No. 3.6e-57;
Matches 164; Conservative 130; Mismatches 249; Indels 74; Gaps 11;

QY      3 HKDLQRLIDIGEDNIVAAAHGATRLRLVLTQTKVDVROSDDDDPKGTFTGGMFO 62
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      24 YELARKIVENGVKENINGLTHCTTRLRFLKNEKANDKELKMDGIVTMRAGGYQ 83
QY      63 IIVGPDVDHFKELDDA--TSKDIADVSTEQLDVYVANNANWFSRAVKYADI FVPLPIL 120
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      84 VVIG-NHVVVEEYVYKAGNLTFDEAVSTKKRP-----FDMIDISGCFPLAI 134
QY      121 LVGGGLMANVNVVAODLFGPQSLVEMFPQISGVAEMINLMASAPAFPLVVGFTATRK 180
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      135 LAAGMIRGLTFLV-----FLGAFDRSGGFVFNVDNIGDSVFQFMPVIGLTAAR 185
QY      181 RFGNEFLGAGIGMANVFPVLVNGYDVAAITMTAGEMPMSLFGLDVAQA----- 229
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      186 KKVNEFVGMILGALMNPSL-----SLELSGAABAPLTITISGTFEAPITYQVFGIW 241
QY      230 ---GYQTVLPVLVSWILATIEKFLHKLMTADFLITPVLTLLTGFLTFIAGPAMR 286
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      242 IARNVASSVPIIFIVLASQIQKPLKLVPEMINAFVFPFTVLITWPLGLVGPVFT 301
QY      287 WTGDLIHLGLQGYDGPVGGLLFGLVSPVITVGHOSFPIEELF-NOGGSFTFAT 345
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      302 FATDILMAGFETILASPVIYGAIVGFQWQILVMFGLHVAIYPMGIMQPSVGMONIMTP 361
QY      346 ASMANIAQACLAVERFLAKSEKLGAGAGSVSAVLGTEPAIFGVNLRMPFIYIG 405
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      362 VAVVSFGTALPTALPYFKLRNPKDPAIAIPAVSGIVGTEPAIFGTLPRKKIIFLTGV 421
QY      406 TAAIGGALLALPDIKAVALGAGFLGVSI-----DAPDMWFLVCAVTFPIARGAALA 460
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      422 GGALGAGVSGLMNLSMNGGGLGIFTIPYIARPDGDLTDVINVLGIALAMVVSFLTFE 481
QY      461 YGLVYVRNGSIDPDATAPVPAGTTKAEAEAPAFSNDST--TIQAPLTGALISVS 518
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      482 F-----WKDEAGETDIOKKSGEIKYKPTIOGIAPLANAK 517
QY      519 DAMPASGKLGSGVALVPTKGLVSPVSGKIIVAFPSGNAFAVTRTADGSNDVIMHIF 578
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      518 DAAFAQGLTGRIILYPEKGEVARAPDGTWMLFTKALIGVSR---TGLELHVL 573
QY      579 DTVNINGTHFNPLKQOG 595
DB      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      574 DTVOLEBKRYFESLVOQG 590

```

Search completed: March 7, 2005, 21:57:09
Job time : 46 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 21:54:07 ; Search time 136 Seconds

(Without alignments)
1599.255 Million cell updates/sec

Title: US-10-019-284B-2

Perfect score: 3342
Sequence: 1 MDHDLAQRIRLDIGEDNI.....TEAGNLLNVAKKEAVPATP 661

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1391452 seqs, 329044822 residues

Total number of hits satisfying chosen parameters: 1391452

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/ECTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3310	99.0	661	9	US-09-738-626-6404
2	2332	69.8	468	15	US-10-450-055-10
3	957	28.6	617	15	US-10-282-122A-60875
4	956	28.6	627	9	US-09-815-242-13467
5	949	28.4	655	16	US-10-474-776-375
6	949	28.4	655	17	US-10-474-928-3912
7	801.5	24.0	612	15	US-10-282-122A-73843
8	799.5	23.9	612	16	US-10-474-776-253
9	799.5	23.9	612	17	US-10-472-928-1038
10	785.5	23.5	620	15	US-10-282-122A-74439
11	738	22.1	379	15	US-10-282-122A-70209
12	725.5	21.7	683	9	US-09-738-626-6961
13	725.5	21.7	683	15	US-10-450-055-26

14	627	18.8	451	15	US-10-282-122A-57628	Sequence 57628, A
15	584	17.5	381	16	US-10-474-776-367	Sequence 367, App
16	554	16.6	453	15	US-10-282-122A-75796	Sequence 75796, A
17	550	16.5	484	9	US-09-815-242-10809	Sequence 10809, A
18	550	16.5	484	15	US-10-282-122A-56917	Sequence 56917, A
19	548.5	16.4	429	15	US-10-282-122A-73368	Sequence 73368, A
20	538	16.1	455	9	US-09-815-242-13794	Sequence 13794, A
21	531	15.9	453	15	US-10-282-122A-55853	Sequence 55853, A
22	529	15.8	334	17	US-10-472-928-3552	Sequence 3552, App
23	522	15.6	484	9	US-09-815-242-12272	Sequence 12272, A
24	517	15.5	484	15	US-10-282-122A-44138	Sequence 44138, A
25	517	15.5	484	17	US-10-470-0488-229	Sequence 229, App
26	511	15.3	474	15	US-10-282-122A-70588	Sequence 70588, A
27	510	15.3	454	15	US-10-282-122A-45789	Sequence 45789, A
28	510	15.3	454	15	US-10-282-122A-68601	Sequence 68601, A
29	507	15.2	466	15	US-10-282-122A-53222	Sequence 53222, A
30	496	14.8	526	15	US-10-282-122A-57960	Sequence 57960, A
31	495.5	14.8	450	15	US-10-282-122A-71596	Sequence 71596, A
32	490.5	14.7	439	9	US-09-815-242-5660	Sequence 5660, App
33	481	14.4	403	15	US-10-282-122A-52840	Sequence 52840, A
34	472	14.1	545	14	US-10-156-761-11459	Sequence 11459, A
35	464	13.9	454	15	US-10-282-122A-45636	Sequence 45636, A
36	462.5	13.8	480	15	US-10-282-122A-60159	Sequence 60159, A
37	450.5	13.5	476	15	US-10-282-122A-67341	Sequence 67341, A
38	438	13.1	474	9	US-09-815-242-10243	Sequence 10243, A
39	438	13.1	474	15	US-10-282-122A-56632	Sequence 56632, A
40	437	13.1	482	15	US-10-282-122A-76383	Sequence 76383, A
41	436.5	13.1	485	15	US-10-282-122A-43101	Sequence 43101, A
42	434	13.0	483	15	US-10-282-122A-56083	Sequence 56083, A
43	418.5	12.5	477	15	US-10-282-122A-60299	Sequence 60299, A
44	374	11.2	249	16	US-10-474-776-368	Sequence 368, App
45	374	11.2	249	17	US-10-472-928-3554	Sequence 3554, App

ALIGNMENTS

RESULT 1
US-09-738-626-6404
Sequence 6404, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6404
LENGTH: 661
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6404
Query Match 99.0%; Score 3310; DB 9; Length 661;
Best Local Similarity 98.9%; Pred. No. 2.1e-27;
Matches 654; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```

QY      1 MDHKLQALRIKIDIGEDNIYAAACATRLRLVLTQKQVDQFQSDDDPDLKGTETGGM 60
Db      1 MDHKLQALRIKIDIGEDNIYAAACATRLRLVLTQKQVDQFQSDDDPDLKGTETGGM 60
QY      61 FOIIVPGDVHVFKEILDATSKDIAVSTEOIKDVANNANMFESRAVYVLDIFVPLPI 120
Db      61 FOIIVPGDVHVFKEILDATSKDIAVSTEOIKDVANNANMFESRAVYVLDIFVPLPI 120
QY      121 LVGGGLMAINNVLVAODLFGPQSVEMEPQISGVAMINIMASAPFAFLPVLVGFTATK 180
Db      121 LVGGGLMAINNVLVAODLFGPQSVEMEPQISGVAMINIMASAPFAFLPVLVGFTATK 180
QY      121 LVGGGLMAINNVLVAODLFGPQSVEMEPQISGVAMINIMASAPFAFLPVLVGFTATK 180
Db      121 LVGGGLMAINNVLVAODLFGPQSVEMEPQISGVAMINIMASAPFAFLPVLVGFTATK 180
QY      181 RFGGNEFGAGIGAMVFPPLVNGVDVAATMAAGMPMWSLFGLDVAAGQGTPLPVLY 240
Db      181 RFGGNEFGAGIGAMVFPPLVNGVDVAATMAAGMPMWSLFGLDVAAGQGTPLPVLY 240
QY      241 VSMILATIEKFLHKLMTGADFLIPVLTLLTGFTLTAIGPAMRWGDLTAHGLQGLY 300
Db      241 VSMILATIEKFLHKLMTGADFLIPVLTLLTGFTLTAIGPAMRWGDLTAHGLQGLY 300
QY      301 DFGGVGGLLFGLVSPVITGTHQSFPPIELELFNQGSSFIPTASMANIAQGAACIAY 360
Db      301 DFGGVGGLLFGLVSPVITGTHQSFPPIELELFNQGSSFIPTASMANIAQGAACIAY 360
QY      361 FFLASEKTLKGLAGASGVSAVLGTEPAIFGVNLFRLRPFYIGITAAIGGALLPDIK 420
Db      361 FFLASEKTLKGLAGASGVSAVLGTEPAIFGVNLFRLRPFYIGITAAIGGALLPDIK 420
QY      421 AVALAAGFLGVSDAPDMWFLVCAVTEFVIAFGAIAVGLYVRNGSIDDPATAP 480
Db      421 AVALAAGFLGVSDAPDMWFLVCAVTEFVIAFGAIAVGLYVRNGSIDDPATAP 480
QY      481 VPAGTTKAEAEAPAFSSNDSTIIQAPLTGEAIALSVSDAMFASGKLGSVAIVPTKQL 540
Db      481 VPAGTTKAEAEAPAFSSNDSTIIQAPLTGEAIALSVSDAMFASGKLGSVAIVPTKQL 540
QY      541 VSPVSGKIVVAFPSGHAFAVTRTKADSGNVDILMHIGFDVNLNTHNPPLKKQDEKVA 600
Db      541 VSPVSGKIVVAFPSGHAFAVTRTKADSGNVDILMHIGFDVNLNTHNPPLKKQDEKVA 600
QY      601 GELLCEFDIDAIIKAGYEVTTPVIVSNYKKTGPVNTYGLGEIEAGANILNVAKKEAVPAT 660
Db      601 GELLCEFDIDAIIKAGYEVTTPVIVSNYKKTGPVNTYGLGEIEAGANILNVAKKEAVPAT 660
QY      661 P 661
Db      661 P 661

RESULT 2
US-10-450-055-10
; Sequence 10, Application US/10450055
; Publication No. US20040043953A1
; GENERAL INFORMATION:
; APPLICANT: BASF Aktiengesellschaft
; TITLE OF INVENTION: No. US20040043953A1el genes of Corynebacterium
; FILE REFERENCE: 936 2000
; CURRENT APPLICATION NUMBER: US/10/450,055
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PatentIn Vers. 2.0
; SEQ ID NO 10
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-10-450-055-10

Query Match          69.8%; Score 2332; DB 15; Length 468;
Best Local Similarity 98.5%; Pred. No. 36-192;
Matches 461; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY      194 MAMVFPPLVNGYDAATWTAGEMPWMSLFGLDVAAGQGTPLPVLVSMILATIEKFLH 253

```

```

Db      1 MAMVFPPLVNGYDAATWTAGEMPWMSLFGLDVAAGQGTPLPVLVSMILATIEKFLH 60
QY      254 KRLMGADFLITTPVLTLLTGFTLTAIGPAMRWGDLTAHGLQGLYDFGSPVGGILFGL 313
Db      61 KRLMGADFLITTPVLTLLTGFTLTAIGPAMRWGDLTAHGLQGLYDFGSPVGGILFGL 120
QY      314 VYSPVITGTHQSFPPIELELFNQGSSFIPTASMANIAQGAACIAYVFLAKSEKLGIA 373
Db      121 VYSPVITGTHQSFPPIELELFNQGSSFIPTASMANIAQGAACIAYVFLAKSEKLGIA 180
QY      374 GASGSAYLIGITEPAIFGVNLFRLRMPFYIGITAAIGGALLPDIKAVLGAAGFLGV 433
Db      181 GASGSAYLIGITEPAIFGVNLFRLRMPFYIGITAAIGGALLPDIKAVLGAAGFLGV 240
QY      434 SIDAPDMWFLVCAVTEFVIAFGAIAVGLYVRNGSIDDPATAPVPAGTTKAEAP 493
Db      241 SIDAPDMWFLVCAVTEFVIAFGAIAVGLYVRNGSIDDPATAPVPAGTTKAEAP 300
QY      494 AEFNSDSTIIQAPLTGEAIALSVSDAMFASGKLGSVAIVPTKQLVSPVSGKIVVAF 553
Db      301 AEFNSDSTIIQAPLTGEAIALSVSDAMFASGKLGSVAIVPTKQLVSPVSGKIVVAF 360
QY      554 SGHAFVTRTKAEDSGNVDILMHIGFDVNLNTHNPPLKKQDEKVAKELLCEFDIDAIIK 613
Db      361 SGHAFVTRTKAEDSGNVDILMHIGFDVNLNTHNPPLKKQDEKVAKELLCEFDIDAIIK 420
QY      614 AAGYEVTTPVIVSNYKKTGPVNTYGLGEIEAGANILNVAKKEAVPATP 661
Db      421 AAGYEVTTPVIVSNYKKTGPVNTYGLGEIEAGANILNVAKKEAVPATP 468

RESULT 3
US-10-282-122A-60875
; Sequence 60875, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16

```

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 60875
; LENGTH: 617
; TYPE: PRT
; ORGANISM: Listeria monocytogenes
US-10-282-122A-60875

Query Match      28.6%; Score 957; DB 15; Length 617;
Best Local Similarity 35.1%; Pred. No. 1.8e-73;
Matches 225; Conservative 124; Mismatches 228; Indels 64; Gaps 17;

QY 1 MDHKLORILLDIGEDNIVAAAHCAATRLRLVLTQTKVDQSDDDDLKGFETGGM 60
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MDYQIAKEILANVGSEENVRSVHCATRLRFLVKKERADKQIISIGSVENNAQ 60

QY 61 FQIIIVG--PGDVHV---FKEILDATSKDIAVSTEQDKDVANNANWFSSRAVKVLADIFV 115
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 LQVIIGNTVGDVYKALGFTKLTDGDSIAKGT---KD---SDGNFLSKAIDVISGIFT 114

QY 116 PLIPILVGGGLMAINNVLAQDLFGPQSLVEMFQISGVAEMINIMASAPAFPLVAVG 175
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 115 PILGALAGGGMKGLMILTT---FG-----WLTSSTGYOILYAADSVFVFLPLILA 165

QY 176 FFAATRFQGNFELGAGIGAMVFPPLVNGVDVAATMTAGEMMWSLFGLDVAQAGYQSTV 235
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 166 YPAARFEGNPPVALAAGALVPTMINIFNENCAHTFLQIP-----VLMSTSFVS 217

QY 236 LPLVNSMILATIEKFLHKLMTADFLITPVLTLLTGFLFPIAIGPAMRWGDLAAG 295
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 218 IPIILAVFLSLIERLNKIHAAATPLTPMCLMLVPLTFLAAGPLGTPISGLAAG 277

QY 236 LQGLYDFGPGVGGLLFGLVYSPVITGLHQSPPPIEL-ELFNQGSFIFATASMANIAG 354
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 278 YPIFYNLSPIVAGAFGAFWQVLVIFGIHWGFVPIMINLSRGYRDTMIAMGPSNFAQA 337

QY 355 AACLAVFLAKSEKLGKLAGASGVSALVGTTERPAIGVNLRLMPPYITIGTAAIGAL 414
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 338 GMSLGFELTKRPENVALIAGSALATGFGITESPISYGLTKPKPVALISAGATGALV 397

QY 415 ALFDIKAVALGAAGFLGVNSIDAPDWV---MFLVCAVVTFFVIAFGAIAAY-----GLYLV 466
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 398 -----GAAGSSGANA-IPGLITPIPIGKPFVGFIL--GIAVAYIIISAIGTYF 444

QY 467 RRRGSDPDATAAPVPAGTTKAABAPAFSNDSTIIQAPLTGEAIALSSVSDAMFASGK 526
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 445 ---GKDEMAADGI---APTTEAKETGVE---AEVIVSPIRGNIVPLNEVDEAFSAGL 494

QY 527 LGSVAIVPTKQOLVSPVSGKIIVAPFSGHAFAVRTKAEDGSNDIIMHIGDPTVNLGT 586
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 495 LKQVAIVPQBEKLTISPVGITETAPPTGHAIGIRS---DKGVELIHVGDFVQLNCK 550

QY 587 HENPLKKGDEVKAGELCEPDIDAKAGYEVTTPVIVSN 627
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 551 YFKLVAQSDRVLVGQALLFEDLEAIKAGYDITTPIVTN 591

RESULT 4
US-09-815-242-13467
; Sequence 13467, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haseibeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; Prokaryotes
```

```
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 13467
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13467

Query Match      28.6%; Score 956; DB 9; Length 627;
Best Local Similarity 33.3%; Pred. No. 2.3e-73;
Matches 222; Conservative 141; Mismatches 246; Indels 58; Gaps 14;

QY 1 MDHKLORILLDIGEDNIVAAAHCAATRLRLVLTQTKVDQSDDDDLKGFETGGM 60
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MNNQSIACKVIDALGRENVSVAHCATRLRMVDEBEKINKEVENLEKQAGAFNSQ 60

QY 61 FQIIIVGPDVHVEK---LDDATSKDIAVSTEQDKDVANNANWFSSRAVKVLADI FVPL 117
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 YQIIFGTGVNKMVDEVVVLGLPTS-----SKDDKAEVAKQGNFQRAIRTFGVFVFI 115

QY 118 IPIILVGGGLMAINNVLAQDLFGPQSLVEMFQISGVAEMINIMASAPAFPLVLYGT 177
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 116 IPVIAVATGIVKGVNRLFNALNEM-----PLRDFEYTYQIILTDFAFIILPGLVNS 165

QY 178 ATKRFQGNFELGAGIGAMVFPPLVNGVDVAATMTAGEMMWSLFGLDVAQAGYQSTV 237
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 166 TRRVFGNPNAGVITGMMLVSGSLPNANAVA---OGGEVTAMNFGF-IPVGLQGSVLP 221

QY 238 VLVNSMILATIEKFLHKLMTADFLITPVLTLLTGFLFPIAIGPAMRWGDLAAGIQ 297
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 222 ARIIGVAKKEKAVRKVVPDVITDLVTPVTLVMSITLGVLGVFVHVENVILIAVK 281

QY 238 GLYDFGPGVGGLLFGLVYSPVITGLHQSPPPIELFNQGSFIF-ATASMANIAGQA 356
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 282 AILSNPFGIGFLIGVQLVYSGVHHIFMLEVQLLAADANFNNIITAAATACGA 341

QY 357 CLAIVFLAKSEKLGKLAGASGVSALVGTTERPAIGVNLRLMPPYITIGTAAIGAL 416
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 342 TVAVGKTKNPKLTALPAALSAFLGITEPAIFVNIRFKRPFLSLIAGIAGGLAST 401

QY 417 FDIKAVALGAAGFLGVNSIDAPDWV-----MFLVCAVVTFFVIAFGAIAAYLYVR 468
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 402 -----LGLATGNGIITIPGTMVYNGQLPQVLLVAAVSAFALGALTVMFGV----- 449

QY 469 NSGIDPDATAAPVPAGTTKAABE-APAFSNDSTIIQAPLTGEAIALSSVSDAMFASGKL 527
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 450 --EDEVDAITAAKQAEVABEKEEVAPALQNETLV--TPVIGDVVALLDVNDPVSSGAM 505

QY 528 GSGVAIVPTKQOLVSPVSGKIIVAPFSGHAFAVRTKAEDGSNDIIMHIGDPTVNLGNH 587
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 506 GGGIAVFPQGVVYVLADAEVSIAPPTGHAIGLTKR---NGAEVLHVGIDTVSMNGDG 561

QY 588 FNPILKKGDEVKAGELCEPDIDAKAGYEVTTPVIVSN---YKKTGPVNTYGGAEIEA 644
   |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 563 PEAKVAQGNKVYAGGVLTFTPSNKTIAAGGLDOTTMTVITNTADVASVAPVAT---GSYSK 618
```

OY	645	GANLINV	651
		::	
Db	619	GDVIEV	625

RESULT 5
US-10-474-776-375

```

: Sequence 375: Application US/10474776
: Publication No. US20040110181A1
: GENERAL INFORMATION:
:   APPLICANT: Wyeth
:   TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCODING POLYPEPTIDES
:   TITLE OF INVENTION: ANTIGENS AND USES THEREOF
:   FILE REFERENCE: AM100649-PT
:   CURRENT APPLICATION NUMBER: US/10/474,776
:   CURRENT FILING DATE: 2003-10-14
:   NUMBER OF SEQ ID NOS: 752
:   SOFTWARE: PatentIn version 3.1
:   SEQ ID NO 375
:   LENGTH: 655
:   TYPE: PRN
:   ORGANISM: Streptococcus pneumoniae
:   US-10-474-776-375

```

Query Match	28.4%	Score 949;	DB 16;	Length 655;
Best Local Similarity	34.9%	Pred. No. 9.8e-73;		
Matches 234;	Conservative 125;	Mismatches 253;	Indels 56;	Gaps 19

```

OY 7 NORILFIDGSDNIVAAAHNCAIRLELVKQTRDVROSLDDPDLKGFTEFGAFQIIVG 66
OY 8 AKDILQAGKKNENTVAIVHCAIRKRFVLGDCKKANVKAIESIPAVKGFITNAAGFOYIIG 67
OY 67 PGVDVHFKEILDATSKDIAVSTEQDKOVANNANWESRAVKIADIIVPLIPLIVGGGL 126
Db 68 -NDVPIFEND- TAVSGIEGVSKAKSAKSNQVNVGVWTTTAAIEFTPIIPALIYGGI 125
OY 127 LMAINNVL- -YAODLFGPSOLVEMEPQJSGVAEMINLMASAPFAFPLVLFETATKRRGG 184
Db 126 ILGRNRVLEGVHWSKLDGKITTSSQFWAGNVHFIEMLGEALIFQFLPVGIIWASSRKKG 185
OY 185 NEFLAGIGMAVPEPTLVNGYDVAAITMAGEMP- -MNSLFGIDVAOAGQSTVLPRVLVS 242
Db 186 SQIIGIYVGI CLVSPQILNAYAVASTPRAADIAAMVMWFGFYFNIRIGYQAOVIPALLAG 245
OY 243 WILATIEKELHKLMTGADDELITPVLITLLTGFLLIAGPAMRWGBLHAGIQLYDF 302
Db 246 LSLSTLELFWMKHILREVISMIFVPEPLSLIPALLIANTVLGP- -IGWTIQOGLSSVLA 301
OY 303 G- -GPVG- -LLFLGLVSPVITVITGLHOSFPPIELELF- -NQGSGFIFETASMANIAGOA 356
Db 302 GLTGVKMLPFAIRGALYAPFVITGLHMTNAIDQLADAGGTALWPMIALMSINAGOSA 361
OY 357 CLAVFELAK- -SEKLGLAGASGVSAVIGITEPARIFGVNLRLRMPPIYIGITGALIGALIA 415
Db 362 VFAYYFMRHBEREAOVSIPATISAYIGTEPAPLFGVANKIYFVAGMTSALAGMISV 421
OY 416 LEDIKAVLAAAGFLGVVSIDAPDWVF- - -LCAVVTFYIAGCAALAYGLYVRNGS 471
Db 422 TFWNTAASIGIGLPGIILSIOPQYMLPAGIMLVAIVPMILTF- - - - -EPRKAGL 472
OY 472 IDPDATAVPBAGTTKABE- - -APAEFSNDSTI- - - - -IQALTEGALALSVSUDAM 521
Db 473 FTK- - - - -TEGDNLIOAEFVAOEAEFVNHBPELTSEVILISPLTGOVKELSQATPI 525
OY 522 FASGLGSGVAVIPRKQOLVSPVSGKIYVAPPSGAFAVRKAREGSDVNDLIMHIGPTV 581
Db 526 FASGVMGGVAVIEBQGLISIPVNGSTVVLPEPTKALIGI- - -VSDBG- -VELLHIGMDTV 581
OY 582 NUNGTNHPMLKQGBEVAGELLECFDIDAIKAAGYEYTTETPIVSN- - -YKKTGPVNTYG 638
Db 582 GLDGKGFESLVVQGBHTVVGQQLIRFMDMVIKKAAGLVETETVIIITNOAYATATIP- -GYTP 640
OY 639 LGEBEAGANLL 649

```

Db 641 T-TIQAASIM 650

RESULT 6
US-10-472-928-3912

```

Sequence 3912, Application US/10472928
Publication No. US20050020813A1
GENERAL INFORMATION:
APPLICANT: CHIRON SPA
APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
FILE REFERENCE: P026926WO
CURRENT APPLICATION NUMBER: US/10/472,928
CURRENT FILING DATE: 2003-09-26
PRIOR APPLICATION NUMBER: GB-0107658.7
PRIOR FILING DATE: 2001-03-27
NUMBER OF SEQ ID NOS: 4979
SOFTWARE: SeqMan9, version 1.03
SEQ ID NO 3912
LENGTH: 655
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
FEATURE:
OTHER INFORMATION: PTS system, IYAC components
OTHER INFORMATION: Cellular location: membrane
OTHER INFORMATION: Similar to strain R6 sequence 15903741 (0.E+01)
US-10-472-928-3912

```

Query Match	28.4%;	Score 949;	DB 17;	Length 655;
Best Local Similarity	34.9%;	Pred. No. 9.8e-73;		
Matches 234;	Conservative 128;	Mismatches 253;	Indels 56;	Gaps 19;

```

OY      AOCILDDIGGEENIYAAAHACATRLTYLVCOKRDVROSLDDPRLCKTFENGQFQIY 66
Db      AKOLLALGEGKENVATVTHCATRKRFPVLDGDKKAVATISBIPVKGFTFMAGQFOYIIG 67
OY      67 PGVDHAFVELD DATSKDI AVSTEOCLKDVANNANMFESRAKVLADI FVPLIPLVGGGL 126
Db      68 -NDVPIFYNDP--TAVSGIEGVSKEAAKSAKSNQVVGWVTLAEIFTPILIPALVIGGL 125
OY      127 LMAINNVL--VAODIFGPOSIVEMFPQISGVAENIMLMAKAPFELPYLVGFATKRFGG 164
Db      126 ILGFRVNVLEGVWSMMDKGTITRESSQFAGVNHFLMFGEBALFOFLRPGITWSVSRKMG 165
OY      185 NEFLGAGIGMAWAFPLVNGVDVATMTMARGEM--MMSLFGLDVAOAGYOCTVLPVLYVS 242
Db      186 SOLGLVILGICVSPQOLNAVA VASTPRADIAANWMMFGYFTYNRIGYQAOYIPALLAG 245
OY      243 WILATTEKELHKRLMGTADFLLTPVLTLLTGLTFILIGPBMWVDDLHAGLOGLYDF 302
Db      246 LLSLYLEIMWHKHIPREVISMIFVPLSLIPALLAHVLDG---IGMTIGQGLSSVLA 301
OY      303 G--GPGV---GLIFGLVTSPIYTTGLHQSFPRIELELF--NQGSGFTRATMSANTIAOGAA 356
Db      302 GLTGPRKMLFGAIFGALYAPFVITGLHMTVAIDTOLADGCTALPMIALSNIAOGSA 361
OY      357 CLAFPELAK--SEKTKSLAGASGVSAVLGITBPALFGVNLRLRMWFFYIGIGTAIGALYIA 415
Db      362 VPAYVYMEHDEBEAOSVLPATISAYLOGETBALRGVNVKITYIPVAGMGSGALAGLUSV 421
OY      416 LPIKIKVALGAAGFLGVNSIDAPDWVF-----LYCAVTVFVIAAGAAIAYGLVLRNGS 471
Db      422 TFWVTAASIGIGGLPEILISIQOYMLPRAGTMLVAIVPMLTF-----FFRAGL 472
OY      472 IDPDATAAPVPACTTAAENE---APAEFSNDSTI-----IQAPLGEALIALSSVDAM 521
Db      473 FTR-----TEGDITMLQAEFVAOBEAEFVNHBEPLTYSVILISPLTGQVKELSQADPI 525
OY      522 FASGKIGSGVALVPTKRGOLVSPVSKIVYAPSGGAFAVRKAKADGASNVLIMHIGFDV 581
Db      526 FASGVWGGGLVLEPSSGELTSPVNGTIVLVEPTKALGCI--VSGEG--VELLHIGMDIV 581

```



```

Db      111 IPAMGGGMLKXVLVLT-----PMLGILQSDQTIAPLTPFGDAPYVLPPLLAYS 161
Qy      178 ATKRFEGNEFLGIGIGMAVFPFLVNGYDVAAATTAGEMPMSLFGLDVAQAGYGTLP 237
      162 ASQKUKVSTSLAMSVAQVLIHPFVQ-----MVQSGNPL-SLFGVPIPAVSQSSVVP 213
Qy      238 VLVVSMILATIEKFLHRLMGTDLPFLTPVLTLLLTGFLTIAIGPAMRWGDDLHAHQ 297
      214 IIMVWLKMYIEKMIAKLTFAVTKSFLOPFLTVLIVSSCIAVLVVGPICVIGEGSLVG 273
Qy      298 GLYDFGPGVGLLFGIVYSPVIVTGLHQSFPIELELFNQGSGFPEATASNA----- 349
      274 QYIVDAGWTLTALGAIMPTVMTGMHAFAP-----IFLAASITPDLVLIIP 321
Qy      350 -----NIAOGACIAAFPLAKSEKLGLAGASGVSAVL-GITEPAIFGVNLRAMPFYIG 403
      322 AMLGSLIAGAAASMAVALKSNNNTKOIAFAAGSALLAGITTEPALVGLTKKKPKLV-- 379
Qy      404 ICTAIGALIALF-----DIKAVALGAAGFLGV--VSIDAP-DMWFLVCVATFVIAF 455
      380 --AAMIGGLAGIFAGLTSTKAYLFAVPSLIALPQFIYSDVPNSIVNALIYAVISVITF 437
Db      456 GAALVAGLYLVRRNGSIDPDATAVPVAGTTKAAEAPAESNSTITQAPLTGSAIALS 515
      438 VLAIFYG-----IDEBESS-----SNLEKAVGVSNNKMF-SPISGEIIPUS 478
Qy      516 SVSDAMFASGKLGSGVAIVPTKCOLVSPVSGKIVVAFPSGAFVARTKADGSDNDILMH 575
      479 DVQDKTFSDKLTIGDGAIIIPSEGVAPRPODKTINIFPTGAIGL--KSDEG--VELLIH 534
Qy      576 IGFDIVNLNGHTHNPILKKGDEYKAGELCEFDIDAIIKAGYEVTTPIVSN 627
      535 IGLDFVELKGGGFIHSHEEGDRVFKNQLIFEMDINTLIKTKGYETVTPVIVTN 586
Db      535 IGLDFVELKGGGFIHSHEEGDRVFKNQLIFEMDINTLIKTKGYETVTPVIVTN 586

RESULT 8
US-10-474-776-253
; Sequence 253, Application US/10474776
; Publication No. US20040110181A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: NOVEL STREPTOCOCCUS PNEUMONIAE OPEN READING FRAMES ENCODING POLYNUCLEOTIDE SEQUENCES
; FILE OF INVENTION: ANTIGENS AND USES THEREOF
; FILE REFERENCE: AM100649-PCT
; CURRENT APPLICATION NUMBER: US/10/474,776
; CURRENT FILING DATE: 2003-10-14
; NUMBER OF SEQ ID NOS: 752
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 253
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-474-776-253

Query Match      23.9%; Score 799.5; DB 16; Length 612;
Best Local Similarity 31.4%; Pred. No. 7e-60;
Matches 205; Conservative 119; Mismatches 237; Indels 91; Gaps 18;

Qy      1 MDHKLQRIARDIGSEBNDIVAAAHACATRLRLVLTQKTDVNDROSIDDDPDLKGTETGGM 60
      1 MSYKDTVQKILDIVIGGEKVNVRVTHCVTRLRLLEKIDENLVNDDDVKKI.PGVIGIMKXNQ 60
Db      61 YQIILG-NDVANYIKKFFVKLGNFESDSV-----VOGHKNILEIIRIYAGSMTP 110
      61 FQIIVGPGVDVHFKE---LDDATSKDIAVSTEQKQOVANNANMFSRAVKVLADI.FVPL 117
Qy      118 IFILVGGGLIMAINNVLAQDLFGPOSILVEMFPOISGVAEMINIMASAPAF.LPVLVGT 177
      111 IPAMGGGMLKXVLVLT-----PMLGILQSDQTIAPLTPFGDAPYVLPPLLAYS 161
Db      178 ATKRFEGNEFLGIGIGMAVFPFLVNGYDVAAATTAGEMPMSLFGLDVAQAGYGTLP 237
      162 ASQKUKVSTSLAMSVAQVLIHPFVQ-----MVQSGNPL-SLFGVPIPAVSQSSVVP 213
Qy      238 VLVVSMILATIEKFLHRLMGTDLPFLTPVLTLLLTGFLTIAIGPAMRWGDDLHAHQ 297
      214 IIMVWLKMYIEKMIAKLTFAVTKSFLOPFLTVLIVSSCIAVLVVGPICVIGEGSLVG 273
Qy      298 GLYDFGPGVGLLFGIVYSPVIVTGLHQSFPIELELFNQGSGFPEATASNA----- 349
      274 QYIVDAGWTLTALGAIMPTVMTGMHAFAP-----IFLAASITPDLVLIIP 321
Qy      350 -----NIAOGACIAAFPLAKSEKLGLAGASGVSAVL-GITEPAIFGVNLRAMPFYIG 403
      322 AMLGSLIAGAAASMAVALKSNNNTKOIAFAAGSALLAGITTEPALVGLTKKKPKLV-- 379
Qy      404 ICTAIGALIALF-----DIKAVALGAAGFLGV--VSIDAP-DMWFLVCVATFVIAF 455
      380 --AAMIGGLAGIFAGLTSTKAYLFAVPSLIALPQFIYSDVPNSIVNALIYAVISVITF 437
Db      456 GAALVAGLYLVRRNGSIDPDATAVPVAGTTKAAEAPAESNSTITQAPLTGSAIALS 515
      438 VLAIFYG-----IDEBESS-----SNLEKAVGVSNNKMF-SPISGEIIPUS 478
Qy      516 SVSDAMFASGKLGSGVAIVPTKCOLVSPVSGKIVVAFPSGAFVARTKADGSDNDILMH 575
      479 DVQDKTFSDKLTIGDGAIIIPSEGVAPRPODKTINIFPTGAIGL--KSDEG--VELLIH 534
Qy      576 IGFDIVNLNGHTHNPILKKGDEYKAGELCEFDIDAIIKAGYEVTTPIVSN 627
      535 IGLDFVELKGGGFIHSHEEGDRVFKNQLIFEMDINTLIKTKGYETVTPVIVTN 586

```

```
QY 238 VLVSMILATIEKFLHKLMTADPLITPVLTLLTGLTFLTAIGPMRWGDLAHGQ 297
DB 214 ILLWMLMKYIEKILAKTLTAITSFLOPTLVLSGCIALVWVGPIGVIGEGLSNLVG 273
QY 238 GLYDGPVGGLLFLGLVSPVITGLHOSFPPIELELFNQGSGFFIATASMA----- 349
DB 274 QMYGVAAGMLTLAILGAIIMPFIYMTGMHAFAP-----IFLAASIAITPVLLP 321
QY 330 -----NIAOGAACLAFFLAKSEKLGKLAGASGVSAVL-GITEPAIFGVNLRMPFIYG 403
DB 332 AMLGSNLAQGAASMAVALKSKNNNTKQIAFAAGFSALAGITEPALYGTLLKYPKPLY-- 379
QY 404 IGTAAIGGALTLALF-----DIKAVALGAAGFLGV---VSIDAP-DMWMLVCAVTVFVIAF 455
DB 380 --AAITGGGLAGLFLGSLTSVKAIVFAVPSLLALPOFISDVPSNVALIYAVISVITF 437
QY 456 GAALAYGLYLVRRNSIDPDATAPVPACTTAAEAAPAEFSNDSTIIQAPLTGEAIALS 515
DB 438 VLAYIFG-----IDEESS-----SNLEVAGVSNKKMIF-SPISGEIIPLS 478
QY 516 SVSDAMFASGKLGSGVAIVPTKGQVSPVSGKIIVAFPSGHAFAVRTAEDGSNDVILMH 575
DB 479 DVQDKTFSDKLIIGDVAIIPSEGKYVAFPDGKITNIIPFKHAIGL--KSDEG--VELLIH 534
QY 576 IGFDTVNLNGTHFNPLKQGEVYKAGELICEFDIDAIAKAGYEVTTPVIVSN 627
DB 535 IGLDVELKGGQFISHVEBGDRVFNKQLI FEMDLNLIKTKGYETVTPVIVTN 586
```

RESULT 9

```
US-10-472-928-1038
; Sequence 1038, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926MO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqMan9, version 1.03
; SEQ ID NO: 1038
; LENGTH: 612
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: PTS system, beta-glucosidase-specific ITABC components
; OTHER INFORMATION: Cellular location: membrane
; OTHER INFORMATION: Similar to strain K6 sequence 15902549 (O.E+01)
US-10-472-928-1038
```

Query Match

Best Local Similarity 31.4%; Pred. No. 7e-60; Matches 205; Conservative 119; Mismatches 237; Indels 91; Gaps 18;

```
QY 1 MDKQLAQRILRDIGEDNIVAAHCAITRLIVLKDITDVRQSLDDDDPKGTFETGGM 60
DB 1 MSYKDTQVKILDLVIGEKXVNVTVTCVTRLELDEMLVNDVDVAKIPGVIGIKKKGQ 60
QY 61 FOIITPGDVHVEK-----LDDATSKDIAVSTEQKDVANNANMFSAVKTLADIFVPL 117
DB 61 YQIITG-NDVANYVKEPVYLGNFESDSV-----VQGHGKILIERIETIASMTPI 110
QY 118 IPIIVGGGLMAINNVLAQDLFGPQSLVEMFPQISGVAEMINMASAPFAFLPVLTGFT 177
DB 111 IPAMLGGMKLKLVIL-----PMIGIQSQSQTIAITLTFEGDAPYITLPLLAWS 161
QY 178 ATKREFGNEFLGAGIGMAVVFPTLVNGVDVAATMTAGEMPMSLFGLDVAQAGYGVLP 237
DB 178
```

```
DB 162 ASQKLKVTSTLAMSVAAGVLHLPNFVQ-----MVQSGNPL-SLEGAPVTPASVSSVVP 213
QY 238 VLVSMILATIEKFLHKLMTADPLITPVLTLLTGLTFLTAIGPMRWGDLAHGQ 297
DB 214 ILLWMLMKYIEKILAKTLTAITSFLOPTLVLSGCIALVWVGPIGVIGEGLSNLVG 273
QY 238 GLYDGPVGGLLFLGLVSPVITGLHOSFPPIELELFNQGSGFFIATASMA----- 349
DB 274 QMYGVAAGMLTLAILGAIIMPFIYMTGMHAFAP-----IFLAASIAITPVLLP 321
QY 330 -----NIAOGAACLAFFLAKSEKLGKLAGASGVSAVL-GITEPAIFGVNLRMPFIYG 403
DB 332 AMLGSNLAQGAASMAVALKSKNNNTKQIAFAAGFSALAGITEPALYGTLLKYPKPLY-- 379
QY 404 IGTAAIGGALTLALF-----DIKAVALGAAGFLGV---VSIDAP-DMWMLVCAVTVFVIAF 455
DB 380 --AAITGGGLAGLFLGSLTSVKAIVFAVPSLLALPOFISDVPSNVALIYAVISVITF 437
QY 456 GAALAYGLYLVRRNSIDPDATAPVPACTTAAEAAPAEFSNDSTIIQAPLTGEAIALS 515
DB 438 VLAYIFG-----IDEESS-----SNLEVAGVSNKKMIF-SPISGEIIPLS 478
QY 516 SVSDAMFASGKLGSGVAIVPTKGQVSPVSGKIIVAFPSGHAFAVRTAEDGSNDVILMH 575
DB 479 DVQDKTFSDKLIIGDVAIIPSEGKYVAFPDGKITNIIPFKHAIGL--KSDEG--VELLIH 534
QY 576 IGFDTVNLNGTHFNPLKQGEVYKAGELICEFDIDAIAKAGYEVTTPVIVSN 627
DB 535 IGLDVELKGGQFISHVEBGDRVFNKQLI FEMDLNLIKTKGYETVTPVIVTN 586
```

RESULT 10

```
US-10-282-74439
; Sequence 74439, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liansu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haeselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: EITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See file Wrapper or PALM.
```

NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 74439
 LENGTH: 620
 TYPE: PRF
 ORGANISM: Streptococcus pyogenes
 US-10-282-122A-74439

Query Match 23.5%; Score 785.5; DB 15; Length 620;
 Beet Local Similarity 30.7%; Pred. No. 1.1e-58;
 Matches 195; Conservative 127; Mismatches 263; Indels 51; Gaps 13;

```

QY 1 MDKDLAQRILRDIGEDNIVAAACATRLRLVLTQTKVDQSLDDDDPKLKFETGGM 60
DB 1 MTYQETAKILAAVGGKTNIQVTHCVTLRLVLRKDEKQKQQAIVASIVGWRKNG 60
QY 61 FOIIVPGDVH--VFKELDATSKDIAVSTEQLDVYANNAMFSAVKYLAIDFVPL 117
DB 61 YQIIIG-NDVNNVYQAFSLGHFDNDQDEHSSKAKKSIL-----ERLIETIAGVITPL 112
QY 118 IPIIVGGILMAINNVLAODLFGPQSLVEMFPQISGVAMINIMASAPAPLPUVGRT 177
DB 113 IPIIVGGILMAINNVLAODLFGPQSLVEMFPQISGVAMINIMASAPAPLPUVGRT 177
QY 178 ATKRFQGNFELGAGIGMAVFPVLVNGVDVAATMTAGMPMMSLFGDLVAQGYQGTVP 237
DB 164 AAARFQVTVLTAITIGILHFAV-----AMVAEGKPL-TLFGAPVTPASYSQSVIP 215
QY 238 VLVSMIATIEKFLHKLMTGADFLITPVLTLTLTGFLTLFIAIGPAMRWGDLAAGLQ 297
DB 216 ILMVYLMQYIEKVVRLVPSVWKSFLQPTLITLISGLALVVGPLGVITIGGSLNTML 275
QY 298 GIYDGGPRVGLFLGLVSPVITITGHQSPPIELEFNQGGSPITATMSA-NIAQGA 356
DB 276 ALYHVAWPLALALIGIMPLVMTGMHMAFAPILFASVATDVLILPMLASNLIAQGA 335
QY 357 CLAVEFLAKSEKLGKLAGASGVSAVL-GITEPAIFGVNLRIMPFIIGTALIGGALA 415
DB 336 SLAAVFKTKQKOTROVALAAGISALLAGITEPLVGLTKFKKPLAANAISGGLVGAFLG 395
QY 416 LFDIKAVALLGAAGFLGVNSIDAP---DMVPLVCVVTFVIAFGAALVGLYLRNGS 471
DB 396 FVNINASYTFVPSIIGLPQYINPSGGANFTNALIGTATITVLAFLTWFMG----- 446
QY 472 IDPDATAPVPACTTAKAEAPAFPSNDSTIIQAPITGAILALSSVSDMPASGLGCV 531
DB 447 IDEE---SPKQVSVADMSQVKSGLSTKOT-LYAPWTGEMLFISEVPDETFSKLGGEF 502
QY 532 AIVPTKQGVSPVSGKIVVAFPSGHAFAVRTKAEDSNVDIIMHIGFDVNNNGTHFNPL 591
DB 503 AILPSEGEVYAFPDGEVITFFPTKAAVALK---NTRGVEVLIHGIDIVELKGGQFEOL 558
QY 592 KKQGDVAKAGELLCEFDIDAIKAAGYEVTTPIVNSN 627
DB 559 VSVGDVVRKQALMKMDIDIFITSKGLSPVVVTN 594

```

RESULT 11
 US-10-282-122A-70209
 ; Sequence 70209, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haseibeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zykend, Jwdich
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Foreyth, R.

APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA.034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 PRIOR FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 Remaining Prior Application data removed - See file wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 70209
 LENGTH: 379
 TYPE: PRF
 ORGANISM: Staphylococcus aureus
 US-10-282-122A-70209

Query Match 22.1%; Score 738; DB 15; Length 379;
 Best Local Similarity 39.5%; Pred. No. 6.9e-55;
 Matches 144; Conservative 89; Mismatches 116; Indels 16; Gaps 5;

```

QY 97 ANNANFSAVAVTLADIFVPLIPILVGGGLMAINNVLAODLFGPQSLVEMFPQISGVA 156
DB 1 AOKGNPVRFLITLGDIFIPILPAIVTAGLLMGINNLTLMKGLFQPKALIEVYPOIADIS 60
QY 157 EMINMASAPAFELVUVGFTATKRFQGNFELGAGIGAMVFPVLVNGVDVAATMTAGM 216
DB 61 NIIINVIASATFELPALIGWSSRVFGSPYIGIVGLIMHPOLVSQDLA---KGNV 116
QY 217 PMSLFGDLVADAGYQGTVPVLVSVMIATIEKFLHKLMTGADFLITPVLTLTLTGFL 276
DB 117 PWNLFGLIEKQLANQGVLPVLAAYLAKIEKGLNVHDSIGMLVGPVALLVGTPL 176
QY 277 TFIAGPAMRVGDILLAGLQGLYDPGGPVGGLFGLVYSPVITGHLHOSPPPIELELPN 336
DB 177 AFIIIGPYALLIGTIGTIGTGVTFPIFOHAGMLGAIVGLYAPLVITGLHMFPLAVFQLMG 236
QY 337 Q--GGSFIFATSMANINQGAACLAFFLASEKL---KGLAGASGVSAVGTETPAIFG 391
DB 237 SLSGDTYMPVIAISNICQSSAAGFAMFYVKKRKKVKEBGLJTCISGMLGVTEPAMWG 296
QY 392 VNLRLRMPFYIGIGAAIGALIALFDIKAVALLGAAGFLGV---VSIDAPDMVFLVCAY 448
DB 297 VNLPLKVFPIAIAISICVLGAIVGNNV---LGRVGVGVAFATISIOKEFFPVYLIARA 352
QY 449 VTFVT 453
DB 353 IAIIV 357

```

RESULT 12
 US-09-738-626-6961
 ; Sequence 6961, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI

APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
PRIOR FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: Patent ver. 3.0
SEQ ID NO: 6961
LENGTH: 683
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6961

Query Match 21.7%; Score 725.5; DB 9; Length 683;
Best Local Similarity 30.9%; Pred. No. 2e-53;

Matches 216; Conservative 119; Mismatches 286; Indels 79; Gaps 18;

QY 7 AGRIRIDGEGBNVYAAAHGATRLRLVKTQDVROSLDDPDLKGFETG--GMFOII 64
DB 9 SCHILENIGGPNITSMTHCATRLRFQVKDSIVDQSIDPSVLGVVPGSGTGM-QVY 67
QY 65 VGPQDVHVEKEL-----DQATSKDIAVSTEQDKOVANNAMFSAVKVLDI 113
DB 68 MG-GSVANYQOILKLDGKMFADGEATES-----SSKKEYGVGRKYSWIDYAEFLSDT 122
QY 114 FVPLIPILVGGGLMAINNVVADLFGPQSLVEMFPOISVAMINIMASAPFAPLVL 173
DB 123 FRPILMALGASLITL--LVADTFGLQDFRAPMDEQPTTYVFLSHMWSVFYFLIM 179
QY 174 VGFATKRFGNNEFGAGIGAMVFTLVNGYDVAATATAGEMPWMSLFGLDVAQAGYOG 233
DB 180 VGATARKLGANEMWIGAIIPALTLTP-----EFLASGADTV--TVFGLPMVLNDYSG 231
QY 234 TVLPVLVSWILATIEKFLHKLMTADFLITPVTLTLTGFLFIAIGPMRWGDLA 293
DB 232 QVFPPLIAAIGLWYWEKGLKTIIPBAVQMVFPFSLIMIPATAFLGPRGIGVNGIS 291
QY 294 HGLQGLYDFGPGVGLLFGLVYSPVITGLHQSPPPIELFNO--GGSFIFATASMANIA 352
DB 292 NLEBINNFSPIISIVPLVPLVPLGLHWPPLNAINIQTINTLGYFFIGPMGANMFA 351
QY 353 QGAACLAFLPAKSEKLGLAGAS--GVSAYV--GITEPAIFGNLRLRMPFYIGITPA 408
DB 352 CFGLVTVGFLLSIKERNRANROVSLGGMLAGLIGISEPSLTVGLRKRKTYFRLLPCL 411
QY 409 IGGALIALFDIKAVALGAAGFLGVVSIDAPDMVMEFLVCAVTVFVIAF--GAALAYGLYVR 467
DB 412 AGGIWGFIDIKAVAFVFTSLTITPAMD--PWLGYITIGIAVAFVSMFLVLADY-----R 465
QY 468 RNSGID-----PDAT-AAV-----PAGTTKAEAEPA 494
DB 466 SNEERDEARAKVADKQAEEDLKAEANATPPAPVAAAGAGAGAGAAAGATAVAAPK 525
QY 495 EFSNDSTIIQAPLTGEAIALSVSDAMPASGKLSGVAIVPTKGLVSPVSGKIIVAPPS 554
DB 496 LAAGVVDIVPPLSGKALPLSEVDPPIPAAGKLGEGIALQGTNTVVAADATVILVOKS 585
QY 585 GHAFVTRTKAEDGSNVDLIMHIGFDTVNLNGTHFNPLKKQGEVYAGELLCEFDIDAIA 614
DB 586 GHAFVTRTKAEDGSNVDLIMHIGFDTVNLNGTHFNPLKKQGEVYAGELLCEFDIDAIA 614

DB 586 GHAFVTRTKAEDGSNVDLIMHIGFDTVNLNGTHFNPLKKQGEVYAGELLCEFDIDAIA 614
QY 615 AGYEVTTPVIVSNYKKTGPVNTYGLGETEAGANLINVAKK 654
DB 642 KDLPLTFPVVVSNAKFGIEIGIPADQANSSTVIKVNKG 681

RESULT 13

US-10-450-055-26
Sequence 26, Application US/10450055
Publication No. US20040043953A1
GENERAL INFORMATION:
APPLICANT: BASF Aktiengesellschaft
TITLE OF INVENTION: No. US20040043953A1 genes of Corynebacterium
FILE REFERENCE: 936-2000
CURRENT APPLICATION NUMBER: US/10/450, 055
CURRENT FILING DATE: 2003-06-10
NUMBER OF SEQ ID NOS: 46
SOFTWARE: Patent ver. 2.0
SEQ ID NO: 26
LENGTH: 683
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-10-450-055-26

Query Match 21.7%; Score 725.5; DB 15; Length 683;
Best Local Similarity 30.9%; Pred. No. 2e-53;
Matches 216; Conservative 119; Mismatches 286; Indels 79; Gaps 18;

QY 7 AGRIRIDGEGBNVYAAAHGATRLRLVKTQDVROSLDDPDLKGFETG--GMFOII 64
DB 9 SCHILENIGGPNITSMTHCATRLRFQVKDSIVDQSIDPSVLGVVPGSGTGM-QVY 67
QY 65 VGPQDVHVEKEL-----DQATSKDIAVSTEQDKOVANNAMFSAVKVLDI 113
DB 68 MG-GSVANYQOILKLDGKMFADGEATES-----SSKKEYGVGRKYSWIDYAEFLSDT 122
QY 114 FVPLIPILVGGGLMAINNVVADLFGPQSLVEMFPOISVAMINIMASAPFAPLVL 173
DB 123 FRPILMALGASLITL--LVADTFGLQDFRAPMDEQPTTYVFLSHMWSVFYFLIM 179
QY 174 VGFATKRFGNNEFGAGIGAMVFTLVNGYDVAATATAGEMPWMSLFGLDVAQAGYOG 233
DB 180 VGATARKLGANEMWIGAIIPALTLTP-----EFLASGADTV--TVFGLPMVLNDYSG 231
QY 234 TVLPVLVSWILATIEKFLHKLMTADFLITPVTLTLTGFLFIAIGPMRWGDLA 293
DB 232 QVFPPLIAAIGLWYWEKGLKTIIPBAVQMVFPFSLIMIPATAFLGPRGIGVNGIS 291
QY 294 HGLQGLYDFGPGVGLLFGLVYSPVITGLHQSPPPIELFNO--GGSFIFATASMANIA 352
DB 292 NLEBINNFSPIISIVPLVPLVPLGLHWPPLNAINIQTINTLGYFFIGPMGANMFA 351
QY 353 QGAACLAFLPAKSEKLGLAGAS--GVSAYV--GITEPAIFGNLRLRMPFYIGITPA 408
DB 352 CFGLVTVGFLLSIKERNRANROVSLGGMLAGLIGISEPSLTVGLRKRKTYFRLLPCL 411
QY 409 IGGALIALFDIKAVALGAAGFLGVVSIDAPDMVMEFLVCAVTVFVIAF--GAALAYGLYVR 467
DB 412 AGGIWGFIDIKAVAFVFTSLTITPAMD--PWLGYITIGIAVAFVSMFLVLADY-----R 465
QY 468 RNSGID-----PDAT-AAV-----PAGTTKAEAEPA 494
DB 466 SNEERDEARAKVADKQAEEDLKAEANATPPAPVAAAGAGAGAGAAAGATAVAAPK 525
QY 495 EFSNDSTIIQAPLTGEAIALSVSDAMPASGKLSGVAIVPTKGLVSPVSGKIIVAPPS 554
DB 496 LAAGVVDIVPPLSGKALPLSEVDPPIPAAGKLGEGIALQGTNTVVAADATVILVOKS 585
QY 585 GHAFVTRTKAEDGSNVDLIMHIGFDTVNLNGTHFNPLKKQGEVYAGELLCEFDIDAIA 614
DB 586 GHAFVTRTKAEDGSNVDLIMHIGFDTVNLNGTHFNPLKKQGEVYAGELLCEFDIDAIA 614

Fri Mar 11 12:40:57 2005

us-10-019-284b-2.rapb

Page 10

Search completed: March 7, 2005, 22:08:07
Job time : 139 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 22:05:18 ; Search time 44 Seconds
(without alignments)
1445.438 Million cell updates/sec

Title: US-10-019-284B-2

Perfect score: 3342
Sequence: 1 MDHKLQRIIRDIGEDNI.....ITAGNLANVAKKAVRATP 661

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR 79:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1609.5	48.2	651	2 S44257	phosphotransferase
2	1478.5	44.2	664	1 B32243	phosphotransferase
3	1227	36.7	480	2 S39978	scra protein - Sta
4	1178	35.2	480	2 D90038	PTS system, sucrose
5	1098.5	32.9	627	2 H96951	fusion, PTS system
6	1093	32.7	479	2 F82432	PTS system, sucrose
7	1069	32.0	479	2 JQ0781	sucrose uptake pro
8	1013.5	30.3	470	2 H83926	PTS system, trehal
9	999	29.9	470	2 C69725	phosphotransferase
10	990.5	29.6	632	2 S68599	phosphotransferase
11	960	28.7	627	2 P95200	PTS system IIABC C
12	957	28.6	617	2 AB1167	phosphotransferase
13	956	28.6	627	2 E96067	phosphotransferase
14	949	28.4	655	2 C95220	trehalose PTS syst
15	934	27.9	705	2 A99084	phosphotransferase
16	933.5	27.8	475	2 C89813	hypothetical prote
17	930.5	27.8	630	2 H83686	PTS system, beta-g
18	929.5	27.8	636	2 C83724	PTS system, beta-g
19	927	27.7	633	2 AC1436	PTS system, beta-g
20	924.5	27.7	634	2 AD1078	PTS system, beta-g
21	901	27.0	609	2 I40406	beta-glucoside per
22	896	26.8	631	2 B42603	beta-glucoside spe
23	887	26.5	609	2 T47097	hypothetical prote
24	860.5	25.7	628	2 D97073	PTS system, beta-g
25	859	25.7	617	2 AC1421	beta-glucoside spe
26	829	24.8	618	2 AC1204	phosphotransferase
27	824	24.7	625	2 C25977	phosphotransferase
28	822.5	24.6	636	2 D86807	hypothetical prote
29	816	24.4	456	2 S62331	phosphotransferase

ALIGNMENTS

30	801.5	24.0	612	2	A97935	hypothetical prote
31	799.5	23.9	612	2	B95067	hypothetical prote
32	781	23.4	458	2	H83881	PTS system, sucros
33	770.5	23.1	455	1	WQ8BST	phosphotransferase
34	751.5	22.5	460	2	A39938	phosphotransferase
35	747.5	22.4	640	2	AB1423	beta-glucoside-spe
36	719	21.5	494	2	AG1231	PTS system trehalo
37	706	21.1	459	2	JU0293	levansucrase synth
38	703.5	21.1	494	2	AF1585	PTS system trehalo
39	669.5	20.0	483	2	A10449	protein-Npi-phosph
40	659	19.7	473	2	A98281	trehalose specific
41	655	19.6	473	2	C65236	phosphotransferase
42	652	19.5	473	2	A86122	PTS system enzyme
43	641	19.2	478	2	B82263	PTS system, trehal
44	630.5	18.9	372	2	I39868	sac operon regulat
45	620	18.6	472	2	AF1058	protein-Npi-phosph

ALIGNMENTS

RESULT 1

S44257

phosphotransferase system enzyme II (EC 2.7.1.69), sucrose specific - Pediococcus pentos

C:Species: Pediococcus pentosaceus

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004

C:Accession: S44257

R:Leenhouts, K.K.J.; Bolhuis, A.A.; Kok, J.J.; Venema, G.G.

A:Submitted to the EMBL Data Library, April 1994

A:Description: The sucrose and raffinose operons of *Pediodoccus pentosaceus* PPE1.0.

A:Reference number: S44252

A:Accession: S44257

A:Molecule type: DNA

A:Residues: 1,651 <LEF>

A:Cross-references: UNIPROT: P43470; EMBL: Z32771; NID: g493728; PIDN: CA83668.1; PID: g4755

C:Genetics:

A:Gene: scra

C:Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase

C:Keywords: phosphotransferase

F:488-640/Domain: phosphotransferase system glucose-specific enzyme II, factor III homo

Query Match

Best Local Similarity 48.2%; Score 1609.5; DB 2; Length 651;

Matches 331; Conservative 114; Mismatches 196; Indels 27; Gaps 8;

QY

1

MDHKLQRIIRDIGESDNI

VAAACATRLRLVLTQDVDRQSLDDPDLKGTPTGGM

60

DB

1

MNHQEVADRVNLA

IKNNIQAAACATRLRLVINDSKIQDQALDDADVAGTFTNGQ

59

QY

61

FOIIVPGDVDPVFEKELDDA

ISKDIAVSTEQLKDVVA--NNANMFSRAVKTLADIFVPLI

118

DB

60

YOIIIGPGVDVCKVYDALIV

TKGLK-EVTPDDIKAAVAAGQNKRPIMDFKVSIDFIPPIV

118

QY

119

PLTVGGGLMANNVTVNODL

FGPOSIVEMFQISGVAMINLMASAPAPLPVTVGFTA

178

DB

119

PALVAGGLMANNVTVNODL

FGHFMASVVEVYGLGIMINLMASAPAPLPVTVGFTA

178

QY

179

TKRFGGNEFLGAGIGMAW

FPTLVNNGYDVAATMTAGEMPMSLFGIDVVAQAGYQVTLBPV

238

DB

179

TKRFGGNEFLGAGIGMAW

FPTLVNNGYDVAATMTAGEMPMSLFGIDVVAQAGYQVTLBPV

238

QY

239

LVSNSILATIEKFLHKL

MLMGTDPLITPVTLTLGLFLTAIGPAMRWGLLAHGLG

298

DB

239

LVSNSILATIEKFLHKL

MLMGTDPLITPVTLTLGLFLTAIGPAMRWGLLAHGLG

298

QY

299

LYDFGPGVGLLFGV

VSPIVITGHSQFPTELEF---NOGGSFIPATSMANIQG

354

DB

299

LYDFGPGVGLLFGV

VSPIVITGHSQFPTELEF---NOGGSFIPATSMANIQG

354

QY

355

AACLAFFLAKEKLG

LAGASVSAVLGITPEPAIFGVNLRWPFYIGITGAIGALI

414

DB

355

AACLAFFLAKEKLG

LAGASVSAVLGITPEPAIFGVNLRWPFYIGITGAIGALI

414

QY	415	ALFPIKVVALLGAAGELGVNSIDAPDMMFLVCAVTVFVAFGAALAAVLVIRNNGSIDP	474
Db	419	GLEFVLSVMMSPASVIGFISIFIAKSIRAFEMLSANISFVVAFIPFFIV----AKRTLGGDR	474
QY	475	DATAPVPVAGTTKAEAPAEFNSDSTIIQAPLTGEALIALSSVSADAMPASGKLGSGVAIV	534
Db	475	DQVSPSPATPTSTV-----INVNDEIIISAPVIGASESLKQVNDQVFSAEIMKGGAIV	525
QY	535	PTKQOLVSPVSGKIVVAPPSGHAFAVRPKAEDSGNSVLIIMHIGFDTYNLNGTHNPPLKQ	594
Db	526	PSSQOVAAPADGVTIVTVYDSHRAIGIKTTA----GAEILIHIGLDTYNLNGEHPHTTVQK	581
QY	595	GDEVKAGELLCEFDIDAIKKAQYEVTTPIVSNYKKTGTPVNTYGLGEIEAGANL--LVNA	652
Db	582	GDYHQGGDLGTFDIAALKKANYDPTVWLIVTNTANYANVERLKVTVNQAGEQLVALLTAP	641
QY	653	KKEAVPAT 660	
Db	642	AASSVAAT 649	

RESULT 2

B32243
phosphotransferase system enzyme II (EC 2.7.1.69), sucrose-specific - Streptococcus mutans
C:Species: Streptococcus mutans
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: B32243
R:Sato, Y.; Poy, F.; Jacobson, G.R.; Kuramitsu, H.K.
J. Bacteriol. 171, 263-271, 1989
A:Title: Characterization and sequence analysis of the *scrA* gene encoding enzyme II(*scr*)
A:Reference number: A32243; MUID:89123027; PMID:256656
A:Accession: B32243
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-664 <SAR>
A:Cross-references: UNIPROT:P12655; GB:M22711; NID:G153799; PIDN:AAA26571.1; PID:G153801
C:Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase
C:Keywords: phosphoprotein; phosphotransferase
F:511-664/Domain: phosphotransferase system glucose-specific enzyme II, factor III homol

Query Match	44.2%	Score 1478.5;	DB 1;	Length 664;
Best Local Similarity	46.0%	Pred. NO. 1e-90;		
Matches 311;	Conservative 127;	Mismatches 203;	Indels 35;	Gaps 13;

Db	419	GSMSAAALIGLQVAVASLSGKGFGLPGLSTKASSIPFVVCCELLSPALAAVTVGYC---	4757
Oy	466	VRKNGSIDPDATPAAPPAQTTAKAEAPAEFSNDST-----IIQAPLIGEAIALSSVS	5118
Db	476	--PTKAVDPFAPEAAAEAEIAEEVQ--EIEPEASANKAQVTDEYLAAPLAGEVELTSVN	5322
Oy	519	DAMFASGKLGSQVAIVPTKQOLVSFVSGKIVIAVPPSGHAFAVTRKADGSDVDTLMHIGF	5789
Db	533	DVFSSSEAMNGKIGIAIIPSGNTIYAAEVDGTVIOAIPFTHAYGI--KSPNGA--ELLIHIGI	5888
Oy	579	DTYNLNGTFFHNPILKKQGEDEVKAGEILCFEDDAIKAAEYVTPPIVVSNNYKTPPVNTY-	6378
Db	589	DTYMSMGSKGEQKVQADQDKIKKGDVLGTFDSDKIAEAGLDWTMTFVLTNTADYASVETLA	6448
Oy	638	GLGEIEAGANLLINVAK	653
Db	649	SSGTVAVAGDSLLLEVKK	664

RESULT 3

S39978
scra protein - Staphylococcus xylosus
C|Species: Staphylococcus xylosus
C|Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C|Accession: S39978
R|Magner, E.; Goetz, F.; Brueckner, R.
Mol. Gen. Genet. 241, 33-41, 1993
A|Title: Cloning and characterization of the scra gene encoding the sucrose-specific en
A|Reference number: S39976, MUID:94049686, PMID:8222209
A|Accession: S39978
A|Status: Preliminary
A|Molecule type: DNA
A|Residues: 1-480 <WAG>
A|Superfamily: UNIPROT:P51184, EMBL:X69800, NID:9407905, PIDN:CAA9461.1, PID:94079
C|Superfamily: phosphotransferase system sucrose-specific enzyme II, factor II

RESULT 4
D90038
PTS system, sucrose-specific IIBC component [imported] - *Staphylococcus aureus* (strain N
C:Species: *Staphylococcus aureus*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: D90038
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogunu
me, A.; Mizutani-U, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.: *Shiba, T.*; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratake, K.
Lance 357, 125-1240, 2001
A:Title: Whole genome sequencing of melic111in-resistant *Staphylococcus aureus*.
A:Reference number: A69758; MUID:21311952; PMID:11418146
A:Accession: D90038
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-480 <KUG>
A:Cross-references: UNIPROT:Q99BQ0; GB:BA000018; PID:G13702328; PIDN:BA643469.1; GSPDB:C
A:Experimental source: strain N315
C:Genetics:
A:Gene: acra
C:Superfamily: phosphotransferase system sucrose-specific enzyme II, factor II

Query Match
Beat Local Similarity 35.2%; **Score** 1178; **DB 2;** **Length** 480;
Matches 242; **Conservative** 94; **Mismatches** 130; **Indels** 14; **Gaps** 7;

Query 1 MDHDLARIRLIRDIGGEMIVAAACATRLALVLTQTDVDRQSDDDPDLKGFPTGGM 60
1 MNYQASADILNALGGEBELDMAMHCAIRLRLVINDSELVNEEALNNNDVVGKGFSTGGQ 60
Query 61 FOIIVGPDVDVHFKEKELDATSKDIAVSTEQLDKVANNAMFSRAVVLADIFVPIPI 120
61 YOIIGSGTVKVPSELKLTGKEASTSEYKQASAKMMPLQRFVAMLSDFVPIIPA 119
Query 121 LVGGGLMAINNVLAODLFGPOSIVEMFQISGVAMINLMAAPAPLPVLVGFAT 179
120 IVAGGLMGLNNITLAKDLFSGSKSLIDVYQFAGLAEINIVFANAPFTLLPILIGFSA 179
Query 180 KRFGNEFLGAGIGAMAFPTLVNGYDVAAMTAGEMPMMSLGLDVAQNGYGTULPV 238
180 KRFGNPFPLGAVLGMILVHPSLMSAYDFPKAVEAGKALPYVDVGLMLINQVGGQVLLPM 239
Query 239 LVNSVLIATIRKFLKRLMGTAADFLITPVLTLLTGFLTFPIAGPMKMWGDLVAHGOG 298
240 LVAAYIILASIERKLRKVPYVLNDLITPLLSIFTAFLTFSPVGPITRQLGYNLSGLT 299
Query 299 LVYRGQPVGSLFGLVSPVITVIGHSFPPIELHFL---NQGSGFIPTASMANIAOG 354
300 LYEFQGAIGLIFGLILVAPIVITGMHSHFIVETLLIADATKGTGGSFIPIATMSVAOG 359
Query 355 AAACAVFLAKSEKLKGLAGAGSAVLGTEPRIPFVNRRLMPPIFGITMAIGAL 413
360 GAATAAPFTIKONKLLKGVASAPAGISALLGTEPRMFEVGNKLRPFGALVSGGISAY 419
Query 414 IALFDIKAVLGAAGFLGVSIDA--PDMWFVLCAVVTPIVAFGAALVGLYVRRNGS 471
420 IAFPVKAIALAGTGNLPGFISINPVHAGMLHYFGMTISFI----ALTIVLILSKRAN 475
Db

RESULT 5
H96951
fusion, PTS system, beta-glucosides specific IABC component [imported] - *Clostridium ac*
C:Species: *Clostridium acetobutylicum*
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C:Accession: H96951
R:Noelling, J.; Bieton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium *Cld*
A:Reference number: A69600; MUID:21359325; PMID:21359325
A:Accession: H96951
A:Status: preliminary

A: Molecule type: DNA
A: Residues: 1-627 <NUP>
A: Cross-references: UNIPROT: Q9L8G6; GB: AE001437; PIDN: AAK78403.1; PID: g15023277; GSPDB: G15023277
A: Experimental source: Clostridium acetobutylicum ATCC824
C: GenBank
A: Gene: CAC0423
C: Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase

Query Match 32.9%; Score 1098.5; DB 2; Length 627;
Best Local Similarity 36.8%; Pred. No. 1,8e-65;
Matches 236; Conservative 122; Mismatches 243; Indels 41; Gaps 8;

QY 1 MDHKLQALRILDDGGEDNIVAAAHACATRLRLVKPTKVDVROSLDDDPDLKGTETGGM 60
DB 1 MDYNTAKDILKLLGGKNNMASHACATRLRLVKKEKATKTEKLDGKVGFFSSGQ 60
QY 61 FQIIIVPGDGDVHVFKELDATSKDIAVSTEQDKD---VVANNMVFSAVFLADIFVPL 117
DB 61 YQIIIGGSSVNNVYKAFVFGT---GISSESLSDTKKAANKNNLFFRPAMLSNIFPI 116
QY 118 IPIVLGGGLMAINNVLVAQDLFGFQSLVEMFPQISGVAMINIMASAPFAFLPVLTGFT 177
DB 117 IPAIVASGILMLGLMDLDAFHLVNSK-----SGLYILNMFNAAFQFLPMIAFS 167
QY 178 ATKRECGNEFLGAGIGMAVFPPTLVNGYDVAATMTGEMMMSLFGIDVAQAGQVTLR 237
DB 168 AAREKTEPYLAAAGALMHPDLQNM---TLGEGIKTTINFGNLGNVYQGVTLR 223
QY 238 VLNVEMILATIEKFLKRLMTGADFLITPVLTLLTGLFLFIAIGPAMRWGDLIAHGLQ 297
DB 224 ILISYVMSYIKGLRKIVPEALDILTPLTITMTGFPAMVYVPGGRFVGDELSGLQ 283
QY 298 GLYDEGPGVGLLFGVLVSPVITVGLHQSFPPIELBFGQGG--SFIPTASMANIAQG 354
DB 284 TLYNTTGFSPGVLFGGLYSLVITGTHSHFPAIEAGLLANPAIHKNFLPIWMANVAQG 343
QY 355 AACLAVFPLASKELKGLAGASGVAVLGTTEPAITGVNRLRPFYTGICATAIGGLI 414
DB 344 GALLAVFETPRKPKKKSIAAPASFCLGITTEPAIFGVNLRKYKPFIALGALGAGI 403
QY 415 ALFDIKAVALLGAAGFLGVVSIDAPDMVFLVCAVVFVIAFGAIAAGLYVRNGSIDP 474
DB 404 VFTKAMRAHVGTGIPGLIVKQGSFLNY---IIMILFGCAFITAMVLGITEETEE 459
QY 475 DATAAPVPAITKAENAPAEFSNDSTIIQAPLTGEGALISVSDAMFASGLSGVAIV 534
DB 460 DLNKETVKNDIVEEES-----VSPNGKVVLLKVPDKTFAEGILDDGIGVD 509
QY 535 PTKGOLVSPVSKIVVAFPSGHAFAVRTKAEEDGSNDILMHIGEDTYNLNGTHNPCLKQ 594
DB 510 PEDGEVSPIDDTVHVFETKHAIAKSK---NGEMLIHIGIDTYKMEGNGKSPIND 565
QY 595 GDEVKAGELTCEFDIDALIKAGVEVTPPIVSNKYKTKGPNT 636
DB 566 GBEVKKGDKLIOFDLDLVKEKAVSPVLTITVTHEDMGFVNS 607

RESULT 6
F82432
PRS system, sucrose-specific IIBC component VCA0653 [imported] - Vibrio cholerae (strain C:Jspices: Vibrio cholerae
C:Jspices: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: F82432
R:Heidelberg, J.F.; Blasen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamshayan, V.; Blas, S.; Qin, H.; Dragoti, I.; Sellere, P.
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Accession number: AB2035; MUID:20406833; PMID:10952301
A:Accession: F82432
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-479 <HEI>

A;Cross-references: UNIPROT:Q9KLT8; GB:AE004395; GB:AE003853; MID:9965068; PIDD:AA9653
A;Experimental source: serogroup O1, strain N16961; biotype EI Tor
C;Genetics:
A;Gene: VCA0653
A;Map position: 2
;Superfamily: phosphotransferase system sucrose-specific enzyme II, factor II

Query Match	32.7%	Score 1093;	DB 2;	Length 479;
Best Local Similarity	47.8%;	Pred. No. 3.1e-65;		
Matches 222;	Conservative 85;	Mismatches 151;	Indels 6;	Gaps 3

[illegible]

RESULT 7

sucrose uptake protein - Vibrio alginolyticus
 N/Alternate names: enzyme II-sucrose protein
 C/Species: Vibrio alginolyticus
 C/Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 09-Jul-2004
 C/Accession: J00781
 R/Blatch, G.L.; Scholle, R.R.; Woods, D.R.
 Gene 95, 17-23, 1990
 A/Title: Nucleotide sequence and analysis of the Vibrio alginolyticus sucrose uptake-end
 A/Reference number: J00781; MUID:91071601; PMID:2174811
 A/Accession: J00781
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-479 <BLA>
 A/Cross-references: UNIPROT:P22825; GB:M6766; GB:M30194; NID:g155261; PIDN:AAA27555.1;
 C/Genetic8:
 A/Gene: SCTA
 A/Superfamily: phosphotransferase system sucrose-specific enzyme II, factor II

	Query Match	32.0%	Score 1069;	DB 2;	Length 479;
	Best Local Similarity	46.1%;	Pred. No. 12e-63;		
	Matches	216; Conservative	94;	Mismatches 151;	Indels 8; Gaps 5
OY	1 MDHKLAQRILRDIGGEDNVVAAGCATATRLVLTQDKTVDRDROSLDDPDLKSTFEETGGM	60			
Dd	1 MNYRVAARELLTLLDGSKSNITALLACATLRIVAADEQKIDICQADINLGGVQQFVAQO	60			
OY	61 FOIIYGPDVDAHFVEKLDDATSKDIADVSTEOLKDVVANNNMFRAVKYIADIIVEPLPI	120			

[illegible]

RESULT 8

nba220
 PPS system, trehalose-specific enzyme II, BC component BH2216 [Imported] - *Bacillus halodurans*
 C:Species: *Bacillus halodurans*
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C:Accession: H83926
 R:Takami, H.; Nakabeane, K.; Takaki, Y.; Maeno, G.; Saseki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A>Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A:Reference number: AB3650, WUID:20512582; PMID:11058132
 A:Accession: H83926
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-470 <STO>
 A:Cross-references: UNIPROT:Q9KAS1; GB:AP001514; GB:BA000004; NID:ig10174613; PIDN:BAB059
 A:Experimental source: strain C-125

C:Superfamily: phosphotransferase system sucrose-specific enzyme II, factor II

```
Query Match          30.3%; Score 1013.5; DB 2; Length 470;
Best Local Similarity 44.4%; Pred. No. 5, 8e-60;
Matches 204; Conservative 90; Mismatches 150; Indels 15; Gaps 4

QY 3 HNDLAKRLIRDIQGEEDNTVAAHACAPRLRYLKDKTKVDYROSLLDDPDRLKGFETGGMFQ 62
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 4 YKKEVNAIVEALGGADNIGTATHTCTRLRFVLRDSKVDQEKLESIDYKGSFTNGQIQ 63

QY 63 IIVGPGDVHVEKEI-----DQTSKDIAVSTEQLDKVVANNANMFSAVKVLADI FVP 116
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 64 VIIGGTATDKYKXELVAETGIEAKTKEDV-----KDAAKNTNVFGQAVTLDIIFLP 116

QY 117 LPIIIVGGGLMAINNVLAODLF-GPQSLVEMFPOISGVAEMINIMASAPRAFLPVLVG 175
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 117 ILPAIVTPTAGLLMGINNILLAEIGIPFYDQASVIDIHQOMADFAGIINLIANTAEVFLPGLTG 176

QY 176 FTATRRPGNBEITLGAGIGAMVFPPLTVNGVDVAATMTGEMWMSLFGI.DVAQAGVQTV 235
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 177 WSAVRRFGSESLGLVLGLMVLHPDLNLAMGYGQAQLBESEITPTWNLFGLLITQVQYQGV 236

QY 236 LPLVAVSWLITAEKFLHRLMTGADFLTPVLVTLTLTFLLFIAPGMRWVGDLIAG 295
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 237 LPILSAHLIAKLEIFLRKRPVDSIQQLVNAVALLITQIFAPAIAGPTTFPIIGNITNV 296

QY 296 LQGLVDFGQPVGGLLFGLVSPILVITGLHQSPFIELEFNO-GGSFIFATSMANIAQ 354
```

Db 297 FTSIFAALVGVGLGYLAPLVYGMHTFLAVDLQIGTGFPLVLSIAG 356
Qy 355 AACLAFLAKSEKLGAGASVAVLITTEPAIFGVNLRMPFYIGTAIGALI 414
Db 357 SAALAMFATPBKLGSLSSAVSLGITTEPAMGVNIRYFPVCAIISAIGARI 416
Qy 415 ALPDIKAVALLGAAGFGVVSIDAPDMVMLVCVAVTFVI 453
Db 417 TNGVLANSIGVGGIPGISIQAGFWGVFGIGVIAFIL 455

RESULT 9
C69725
phosphotransferase system enzyme II (EC 2.7.1.69) phosphoenolpyruvate-dependent, trehalo
C/Species: Bacillus subtilis
C/Date: 05-Dec-1997 #sequence, revision 05-Dec-1997 #text_change 09-Jul-2004
C/Accession: C69725; S67929; JCS037; I40497; S67864
R/Kunzt, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beret
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Erttington, J.; Fabre, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall
Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koecker, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maruda, S.; Mausel
Y. M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon,
A/Authors: Schlitch, S.; Schreier, R.; Scoffone, F.; Sekiguchi, J.; Sekowski, A.; Serot
Akeuchi, M.; Tanakaohi, A.; Tanaka, T.; Tepstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winiers, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K
A/Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Darchin, A.
A/Title: The complete genome sequence of the gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: C69725

A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-470 <KUN>
A/Cross-references: UNIPROT:P39794; GB:299108; GB:AL009126; NID:G2633055; PIDN:CAB12609.
A/Experimental source: strain 168
R/Helfert, C.; Gotsche, S.; Dahl, M.K.
Mol. Microbiol. 16, 111-120, 1995
A/Title: Cleavage of trehalose-phosphate in Bacillus subtilis is catalysed by a phospho-
A/Reference number: I40497; MUID:95379486; PMID:7651129
A/Accession: S67929

A/Molecule type: DNA
A/Residues: 324-362, 'L', 364-464, 'G', 466-470 <HEL>
A/Cross-references: EMBL:X80203; NID:G580941
R/Schoeck, F.; Dahl, M.K.
Gene 175, 59-63, 1996
A/Title: Analysis of DNA flanking the trea gene of Bacillus subtilis reveals genes encod
A/Reference number: JCS037; MUID:97074649; PMID:8917076
A/Accession: JCS037

A/Molecule type: DNA
A/Residues: 1-139, 'S', 141-362, 'L', 364-464, 'G', 466-470 <SC2>
A/Cross-references: EMBL:254245; NID:G100450; PIDN:CA91014.1, PID:G100451
C/Comment: This enzyme functions as the specific trehalose transporter. It belongs to th
C/Genetics:
A/Gene: treP, treB
C/Superfamily: phosphotransferase system sucrose-specific enzyme II, factor II
C/Keywords: phosphotransferase; sugar transport system
F/113-133/Domain: transmembrane #status predicted <TM1>
F/160-181/Domain: transmembrane #status predicted <TM2>
F/184-203/Domain: transmembrane #status predicted <TM3>
F/230-245/Domain: transmembrane #status predicted <TM4>
F/263-286/Domain: transmembrane #status predicted <TM5>
F/305-322/Domain: transmembrane #status predicted <TM6>
F/375-399/Domain: transmembrane #status predicted <TM7>
F/402-421/Domain: transmembrane #status predicted <TM8>

Query Match 29.9%; Score 999; DB 2; Length 470;
Best Local Similarity 43.9%; Pred. No. 5-4e-59;
Matches 203; Conservative 93; Mismatches 160; Indels 6; Gaps 4;

Qy 7 AQRIRLDIGGENNYAAAHCAIRLVLKDTQDVROSLDDPDLKGFRTSGMQIYV 66
Db 8 AAOIEAVAGENIAAHTCTVTRFLRFLIDSKVQOEMDQIDVVGKSFSTGQFQVVG 67
Qy 67 PBDVHVFEKLDADSKDAVST-EQLKDVANNANNSRAVKAIVLADIFVPLIPILVGG 125
Db 68 QCTVAKVIAELVKEFG--IGESTKDEVKKASEKNAPLORAVKTLADIFILPLAVTG 125

Qy 126 LMAANNVLAODL-FGPQSLVEMPOISGVAMINLMAAPAFIPVLVGFATKRFPG 184
Db 126 LMGINNLTLAGIFSTKSIQVYVQWADLANMINLLAGTFTPLPILIGSAVRFPG 185

Qy 185 NEFLAAGIGMANVFTLVNGVDVAATMTAGEMPMSTGLDVAQAGQCTVPLVVSRT 244
Db 186 NEFLAIVGVMLVHDLNLMAGYGAABOSGEIPVNNLFGLEVQKVGQGVPLILASYM 245

Qy 245 LATIEKPLHKRLMGADFLITPVLTLTLTGFTLTAIGPAMRWGVDLHAGLQGLYDFG 304
Db 246 LAKIEVFLKRTPEGIQLVAPITLLTGFPASFIITGIFTAIGVLTSGLSVFGSPA 305

Qy 305 PVGGLFGLVSPVITGILHOSFPPIELFNO--GGSFIFATASMANIAQACIAPVF 362
Db 306 ALGGLYGFYSALVITGMHTFLAVDLQILGSKLGGTFMPLMALSINAQSSALAMWF 365

Qy 363 LAKSEKLGAGASVAVLITTEPAIFGVNLRMPFYIGTAIGALIAFDIYAV 422
Db 366 IYKDEKQKGLSTLSTGISAVLITTEPAIFGVNLRMPFYIIMVSSGLAGMYISOGVLAS 425

Qy 423 ALGAAGFGVVSIDAPDMVMLVCVAVTFVIAFGAIVGLY 464
Db 426 SVGVGVPEIRTSIMSQYGAIFGMAIVLIVFAGTAVARF 467

RESULT 10
S68599
phosphotransferase system enzyme II (EC 2.7.1.69), sucrose-specific - Streptococcus sobr
N/Alternate names: sucrose-specific enzyme II
C/Species: Streptococcus sobrinus
A/Variety: strain 6715
C/Date: 14-Feb-1997 #sequence, revision 13-Mar-1997 #text_change 07-May-1999
A/Accession: S68599

R/Chen, Y.Y.M.; Lee, L.N.; LeBlanc, D.J.
Infect. Immun. 61, 2602-2610, 1993
A/Title: Sequence analysis of scvA and scvB from Streptococcus sobrinus 6715.
A/Reference number: S68598; MUID:93273516; PMID:8500898
A/Accession: S68599

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-632 <CH>
A/Cross-references: EMBL:L06791
C/Genetics:
A/Gene: scvA
C/Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase
C/Keywords: phosphotransferase; sugar transport system
F/480-632/Domain: phosphotransferase system glucose-specific enzyme II, factor III homo

Query Match 29.6%; Score 990.5; DB 2; Length 632;
Best Local Similarity 35.5%; Pred. No. 2.8e-57;
Matches 236; Conservative 114; Mismatches 270; Indels 45; Gaps 12;

Qy 1 MDHKLQRIIRLDIGEDNIVAAAHCAIRLVLKDTQDVROSLDDPDLKGFRTSGM 60
Db 1 MDNKOIAKEVIEALGGRDNVRSVAHCAIRLVMVDEAKIKEREMNDKYGAVFNSGQ 60

Qy 61 FOITVGPDVHVFEKLD--ATSKDAVSTEOUKDVANNANNSRAVKAIVLADIFVPL 117
Db 61 VOITFETVYKTIYDEVVDLGLPTS----STGEKQGEAALQGNFQMSRTFGDVFPV 115

Qy 118 IPIIVGGILMAINNVLAODLFGPQSLVEMFPQISGVAEMIN-----LMSAPPAF 169
Db 116 IPVLVATGLFMGRLRLTNDPFLG-----PFGASKQIINNFIYLTQVLTDTARAF 166


```

QY 67 PGVDVHFEKELDDATSKDIASVTEOLKDVANNANWFSRAVKYLAIDFVPLIPILVGGGL 126
DB 68 -NDVPIFYNDF--TAVSGIEGVSKAAKSAKSNQNVGVGWTTLAEFTPIIPALIYGGI 125
QY 127 LMAINNTL--VAODLFGPQSLVEMFPQISGVAKMINMASAPFAFLPVLVGFATKRRFG 184
DB 126 ILGFRNVLBEGVHWSMLDKITTESQFVAGVNHFLMGEALFQPLPVGIGTWSVRKMG 185
QY 185 NEFLGAGIGMAMVPTLVNGYDVAATMTAGEMP--MMSLFGLDVAQAGYQGYVLPVLVS 242
DB 186 SOLIGIVGICLVSPQLNAYASTPAADIAANWVNFGFTVNRIGYQAQVIPALLAG 245
QY 243 WILATIEKFLHKLMTADFLITPVLTLLTGFLTFIAIGPAMRWGDLAHGQGLYDF 302
DB 246 LSLSYLEIFWPKHIEVISMIFVFPFLSLIPALILAHVTLGP---IGMTIGGGLSSVLA 301
QY 303 G--GPVG---GLFGLVYSPVITGLHQSFPPIELFL--NQGSGFIFATSMANIAAGAA 356
DB 302 GLTGPKMLFGAIFGALYAPFVITGLHMTNADITQILDAGGTALMPMLALSNIAQGA 361
QY 357 CLAVFPLAK--SEKLKLAGASGSAVLGITEPAIFGVNLRMPFYIGITAAIGALIA 415
DB 362 VFAVTFMRHDERAKQVSLPATISAVLGTEPALFGVNVKTYFPVAGMTGSLAGMLSV 421
QY 416 LFDIKAVALLGAAGFLGVNSIDAPDMVF---LYCAVTVFVIAFGAALAYGLVVRNGS 471
DB 422 TENVTAASIGIGLPGILSLIOQYMLPFAGTMLVAIVPMLTF-----FFRKA 472
QY 472 IDPPATAAPVPAAGTTKABE---APAEFSNDSTI-----IQAPLTGEALSSVSAM 521
DB 473 FTK-----TEGDTNLQAEFVAQEBAEFVNSHEVELTSVEILSPITQVKELSQATDP 525
QY 522 FASGKLGSVAIVPTKQQLVSPVSGKIYVAPSPGHAFVARTKAEDGSNDVILMHTGPTV 581
DB 526 FASGVWGGLVIEPSSQELTSPVNGTVTLFPTKALIGI--VSDRG--VELLIHIGMDTV 581
QY 582 NLNGTHNPPLKKQGDVYKAGELCEPDLDAIKAGYEVTPIVSNK---YKTKGVPVNYG 638
DB 582 GLDGKGFESLVQGDHVTVGQQLIRFDMVIRKAGLVETETPVITNODAVTATITP-GTYP 640
QY 639 YGLGEIEAGANLL 649
DB 641 T-TIQAGASLM 650

```

RESULT 15

A99084
 phosphotransferase system enzyme II (BC 2.7.1.69), trep [imported] - Streptococcus pneu
 C:Species: Streptococcus pneumoniae
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 C:Accession: A99084
 R:Host: J.A.; Albhorn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
 e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 Y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A>Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A:Reference number: A97872; MUID:21429245; PMID:11544234
 A:Accession: A99084
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1705 <KUR>
 A:Cross-references: UNIPROT:Q8DNI6; GB:AE007317; PIDD:AL00502.1; PID:G15459375; GSPDB:C
 C:Genetics:
 A:Gene: trep
 C:Superfamily: phosphotransferase system enzyme II sucrose-specific; phosphotransferase
 C:Keywords: phosphotransferase

Query Match 27.9%; Score 934; DB 2; Length 705;
 Best Local Similarity 34.2%; Pred. No. 1.9e-54;
 Matches 230; Conservative 130; Mismatches 253; Indels 60; Gaps 18;

```

QY 7 AQRILRDIGGEDNIVAAAHGATRLRLVLTQDKVDROSLDDPDLKGTFFETGMFOIIVG 66
DB 58 AKDLQAIIGKEKENTATVTHCATRRFVLGDDKAAKAIISTAVKGTFTNAQFOVITIG 117
QY 67 PGVDVHFEKELDDATSKDIASVTEOLKDVANNANWFSRAVKYLAIDFVPLIPILVGGGL 126
DB 118 -NDVPIFYNDF--TAVSGIEGVSKAAKSAKSNQNVGVGWTTLAEFTPIIPALIYGGI 175
QY 127 LMAINNTL--VAODLFGPQSLVEMFPQISGVAKMINMASAPFAFLPVLVGFATKRRFG 184
DB 126 ILGFRNVLBEGVHWSMLDKITTESQFVAGVNHFLMGEALFQPLPVGIGTWSVRKMG 235
QY 176 ILGFRNVLBEGVHWSMLDKITTESQFVAGVNHFLMGEALFQPLPVGIGTWSVRKMG 235
DB 185 NEFLGAGIGMAMVPTLVNGYDVAATMTAGEMP--MMSLFGLDVAQAGYQGYVLPVLVS 242
QY 243 WILATIEKFLHKLMTADFLITPVLTLLTGFLTFIAIGPAMRWGDLAHGQGLYDF 302
DB 236 SOLIGIVGICLVSPQLNAYASTPAADIAANWVNFGFTVNRIGYQAQVIPALLAG 295
QY 303 G--GPVG---GLFGLVYSPVITGLHQSFPPIELFL--NQGSGFIFATSMANIAAGAA 356
DB 296 LSLSYLEIFWPKHIEVISMIFVFPFLSLIPALILAHVTLGP---IGMTIGGGLSSVLA 351
QY 357 CLAVFPLAK--SEKLKLAGASGSAVLGITEPAIFGVNLRMPFYIGITAAIGALIA 415
DB 352 GLTGPKMLFGAIFGALYAPFVITGLHMTNADITQILDAGGTALMPMLALSNIAQGA 411
QY 412 VFAVTFMRHDERAKQVSLPATISAVLGTEPALFGVNVKTYFPVAGMTGSLAGMLSV 471
DB 416 LFDIKAVALLGAAGFLGVNSIDAPDMVF---LYCAVTVFVIAFGAALAYGLVVRNGS 471
QY 472 IFNVTAAASIGIGLPGILSLIOQYMLPFAGTMLVAIVPMLTF-----FFRKA 522
DB 472 IDPPATAAPVPAAGTTKABE---APAEFSNDSTI-----IQAPLTGEALSSVSAM 521
QY 523 FTK-----TEGDTNLQAEFVAQEBAEFVNSHEVELTSVEILSPITQVKELSQATDP 575
DB 522 FASGKLGSVAIVPTKQQLVSPVSGKIYVAPSPGHAFVARTKAEDGSNDVILMHTGPTV 581
QY 576 FASGVWGGLVIEPSSQELTSPVNGTVTLFPTKALIGI--VSDRG--VELLIHIGMDTV 631
DB 582 NLNGTHNPPLKKQGDVYKAGELCEPDLDAIKAGYEVTPIVSNK---YKTKGVPVNYG 636
QY 632 GLDGKGFESLVQGDHVTVGQQLIRFDMVIRKAGLVETETPVITNODAVTATITGTYPT 691
DB 637 YGLGEIEAGANLL 649
DB 692 ----TIQAGASLM 700

```

Search completed: March 7, 2005, 22:16:51
 Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 7, 2005, 21:57:13 ; Search time 180 Seconds
(without alignments)
1880,469 Million cell updates/sec

Title: US-10-019-284b-2

Perfect score: 3342

Sequence: 1 MDHKLQRLIRLDIGEDNI.....TEAGANLVAKKEAVPATP 661

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3310	99.0	661	2	Q8NMD6
2	1609.5	48.2	651	1	PTSA_PDPPE
3	1607.5	48.1	651	1	Q88ZV9
4	1508	45.1	650	2	Q9S6S6
5	1478.5	44.2	664	1	PTSA_STPMU
6	1457	43.6	649	2	Q7WMP7
7	1446.5	43.3	647	2	Q74H18
8	1345.5	40.3	643	2	Q82YR5
9	1284	37.8	534	2	Q93ML1
10	1227	36.7	480	1	PTSB_STPMU
11	1186	35.5	480	2	Q8CNB2
12	1183	35.4	480	2	Q8NVT5
13	1183	35.4	480	2	Q6G6U2
14	1178	35.3	480	2	Q6G6U2
15	1178	35.2	480	2	Q99RQ0
16	1178	35.2	480	2	Q7A3V6
17	1160.5	34.7	492	2	Q6SUG9
18	1100	32.9	474	1	PTSB_PSMU
19	1098.5	32.9	474	2	Q9L8G6
20	1093	32.7	479	2	Q9KLT8
21	1069	32.0	479	1	PTSB_VIBAL
22	1013.5	30.3	470	2	Q6SMI3
23	1006	30.1	470	1	PTTB_BACSU
24	999	29.9	470	1	PTTB_BACSU
25	979.5	29.3	480	2	Q8G1N4
26	974	29.1	626	2	Q831B4
27	964.5	28.9	475	2	Q81V69
28	963.5	28.8	475	2	Q6HNI2
29	961.5	28.8	475	2	Q73DL3
30	960	28.7	627	2	Q97PB8
31	957	28.6	617	2	Q8Y904

32	956.5	28.6	475	2	Q63G14
33	956.5	28.6	480	2	Q9F8X3
34	956	28.6	627	2	Q8DMS8
35	954.5	28.6	475	2	Q81H25
36	951.5	28.5	627	2	Q8N2J7
37	949	28.4	655	2	Q97NM9
38	946	28.3	614	2	Q65D36
39	937.5	28.1	627	2	Q8K524
40	935.5	28.0	620	2	Q99Y91
41	934	27.9	619	2	Q65D51
42	934	27.9	705	2	Q8DNI6
43	933.5	27.9	475	2	Q99WC9
44	933.5	27.9	475	2	Q7A7D1
45	933	27.9	674	2	Q8K5K9

ALIGNMENTS

RESULT 1

Q8NMD6 PRELIMINARY; PRT; 661 AA.

AC Q8NMD6; Q6M2J8; 01-OCT-2002 (TREMBLrel. 22, Last created)

DT 01-OCT-2002 (TREMBLrel. 28, Last annotation update)

DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)

DE Phosphotransferase system ITC components, glucose/malose/N-acetylglucosamine-specific (EC 2.7.1.69) (ENZYM II SUCROSE PROTEIN).

GN Name=ptsB; OrderedLocustNames=Cg12642, CG2925;

OS Corynebacterium glutamicum (Brevibacterium flavum).

OC Bacteria; Actinobacteria; Actinobacteriales; Corynebacteriaceae; Corynebacterium.

OX NCBI_TaxID=1718;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025 Nakagawa;

RA Nakagawa S.;

RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";

RL Submitted (MAY-2002) to the EMBL/Genbank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;

RX PubMed=12948626; DOI=10.1016/S0168-1656(03)00154-8;

RA Kalinowski J., Bathe B., Bartels D., Bischoff N., Bott M., Burkovski A., Dusch N., Eggeling L., Eikmanns B.J., Galglat L., Goessmann A., Hartmann M., Huthmacher K., Kramer R., Linke B., McHardy A.C., Meyer F., Moeckel B., Pfeiffer W., Puhler A., Rey D.A., Ruckert C., Rupp O., Sahn H., Wendisch V.F., Wiegand I., Tauch A.;

RT "The complete Corynebacterium glutamicum ATCC 13032 genome sequence and its impact on the production of L-aspartate-derived amino acids and vitamins.";

RT J. Biotechnol. 104:5-25(2003).

DR EMBL: AP005282; BAC00361; --

DR EMBL: BX927155; CAF21304.1; --

DR HSBP; P45618; ZGPR.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0006982; F:protein-N(P)-phosphohistidine-sugar phospho. . .; IEA.

DR GO; GO:0005351; F:sugar porter activity; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.

DR GO; GO:0004401; F:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR GO; GO:0006810; P:transport; IEA.

DR InterPro: IPR011055; DUP_hybrid motif.

DR InterPro: IPR001996; Ptrans_EIIB.

DR InterPro: IPR003352; Ptrans_EIIB.

DR InterPro: IPR001127; Ptrans_EIIB.

DR InterPro: IPR010973; Ptrans_EIIB_sucr.

DR Pfam: PF00358; Ptrans_EIIB_1; 1.

DR Pfam: PF00367; Ptrans_EIIB_1.

DR Pfam: PF02378; Ptrans_EIIB_1.

DR ProDom: PD001476; Ptrans_EIIB_1.

DR ProDom: PD002243; Ptrans_EIIB_1.

TIGRFAMS; TIGR00830; PTB; 1.
 DR TIGRFAMS; TIGR01996; PTS-II-BC-sucr; 1.
 DR PROSITE; PS01035; PTS EIIB CYS; 1.
 KM Complete proteome; Transferrase.
 SQ SEQUENCE 661 AA; 69148 MW; C0E5E1C7833B9F74 CRC64;

Query Match 99.0%; Score 3310; DB 2; Length 661;
 Best Local Similarity 98.9%; Pred. No. 3.3e-191;
 Matches 654; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 MDHKLARILLDIGEGENIVAAACATRLRLVLTQTDVDRQSLDDDDPKGFEETGGM 60
DB 1 MDHKLARILLDIGEGENIVAAACATRLRLVLTQTDVDRQSLDDDDPKGFEETGGM 60
QY 61 FOIIVGPDVHVFEKLDATSKDIASVTEQKDVANANMFSRAVYLAIDFVPLIPI 120
DB 61 FOIIVGPDVHVFEKLDATSKDIASVTEQKDVANANMFSRAVYLAIDFVPLIPI 120
QY 121 LVGGGLMAINNVLAQDLFGPQSLVEMFPQISGVAEMINMASAPFAFLPVLVGFTATK 180
DB 121 LVGGGLMAINNVLAQDLFGPQSLVEMFPQISGVAEMINMASAPFAFLPVLVGFTATK 180
QY 181 REGNEPFGAGIGMAVPEPTLVNGVDVATWTAGEMPMSLFGLDVAQAGYQGYLPLYV 240
DB 181 REGNEPFGAGIGMAVPEPTLVNGVDVATWTAGEMPMSLFGLDVAQAGYQGYLPLYV 240
QY 241 VSMILATIEKFLHKRLMGTADFLITPVLTLLTGFLTFIAGPAMRWGDLIAHGLQGLY 300
DB 241 VSMILATIEKFLHKRLMGTADFLITPVLTLLTGFLTFIAGPAMRWGDLIAHGLQGLY 300
QY 301 DRGGPVGGLLFGVLYSPYITITGLHSGFPRIELETFNQGSGFLPATASMANIQAACLA 360
DB 301 DRGGPVGGLLFGVLYSPYITITGLHSGFPRIELETFNQGSGFLPATASMANIQAACLA 360
QY 361 PFLAKSEKIKGLAGASGAVLIGTEPAIFGVNLRIMFPIGIGTAIGALIALFDIK 420
DB 361 PFLAKSEKIKGLAGASGAVLIGTEPAIFGVNLRIMFPIGIGTAIGALIALFDIK 420
QY 421 AVALGAAGFLGVSDIDAPMVMFLCAVTFVIAFGAIAIYGLVYRNGSIDPDATAAP 480
DB 421 AVALGAAGFLGVSDIDAPMVMFLCAVTFVIAFGAIAIYGLVYRNGSIDPDATAAP 480
QY 481 VPAGTTKARAAAPAESNDSTIIQAFLTEBAIALSSVAMPASGGLSGVAIVPTKGL 540
DB 481 VPAGTTKARAAAPAESNDSTIIQAFLTEBAIALSSVAMPASGGLSGVAIVPTKGL 540
QY 541 VSPVSGKIVAFPSGHAFAVRTKAEDEGSNVLDIMHIGFDTVNLNGTHFNPLKQGEVYA 600
DB 541 VSPVSGKIVAFPSGHAFAVRTKAEDEGSNVLDIMHIGFDTVNLNGTHFNPLKQGEVYA 600
QY 601 GELLCEFDIDAIQAAGYEVTTPIVSNYKKTGPVNTYGGIETAGANILNAKKAIVPAT 660
DB 601 GELLCEFDIDAIQAAGYEVTTPIVSNYKKTGPVNTYGGIETAGANILNAKKAIVPAT 660
QY 661 P 661
DB 661 P 661

```

RESULT 2

PTSA_PDPPE STANDARD; PRT; 651 AA.
 AC P43470;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE PTS system, sucrose-specific IIBAC component (EIIABC-Ser) (Sucrose-
 de permease IIBAC component) (Phosphotransferase enzyme II, ABC
 component) (EC 2.7.1.69) (EII-Ser).
 GN Name=serA;
 OS Pedicoccus pentosaceus.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae; Pedicoccus.
 NCBI_TaxId=1255;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEP1.0;
RA Leenhouts K.K.J., Bolhuis A.A., Kok J.J., Venema G.G.;
RT "The sucrose and raffinose operons of Pedococcus pentosaceus
  PEP1.0";
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: This is a component of the phosphoenolpyruvate-dependent
CC sugar phosphotransferase system (PTS), a major carbohydrate active
CC -transport system. The IID domain contains the sugar binding site
CC and the transmembrane channel; the IIA domain contains the primary
CC phosphoryl group (the donor is phospho-HPr); IIA transfers its
CC the sugar.
CC -1- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
CC histidine + sugar phosphate.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Contains 1 PTS IIA domain.
CC -1- SIMILARITY: Contains 1 PTS EIIB domain.
CC -1- SIMILARITY: Contains 1 PTS EIIC domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; Z22771; CAA83668.1; -
CC EMBL; L32093; AAA25567.1; -
CC PIR; S44257; S44257.
CC HSSP; P20166; 1GPR.
DR InterPro; IPR011055; Dup_hybrid_motif.
DR InterPro; IPR001996; Pctans_EIIB.
DR InterPro; IPR003352; Pctans_EIIC.
DR InterPro; IPR001127; PTS_EIIA.
DR InterPro; IPR010973; PTS_II_BC_sucr.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF00367; PTS_EIIB_1.
DR Pfam; PF02378; PTS_EIIC_1.
DR ProDom; PD001476; Pctans_EIIB; 1.
DR ProDom; PD002243; PTS_EIIA; 1.
DR TIGRFAMS; TIGR00826; EIIB_glc; 1.
DR TIGRFAMS; TIGR00830; PTB; 1.
DR TIGRFAMS; TIGR01995; PTS-II-ABC-beta; 1.
DR TIGRFAMS; TIGR01996; PTS-II-BC-sucr; 1.
DR PROSITE; PS00371; PTS_EIIA_1; 1.
DR PROSITE; PS01035; PTS EIIB CYS; 1.
KW Phosphorylation; Phosphotransferase system; Sugar transport;
  Transferrase; Transmembrane.
FT DOMAIN 1 40 EIIB.
FT DOMAIN 111 418 EIIC.
FT DOMAIN 510 651 EIIA.
FT MOD_RES 25 25 Phosphocysteine (By similarity).
FT MOD_RES 324 324 Phosphohistidine (By similarity).
FT MOD_RES 562 562 Phosphohistidine (By similarity).
SQ SEQUENCE 651 AA; 68454 MW; C07BA09D550A7778 CRC64;

```

Query Match 48.2%; Score 1609.5; DB 1; Length 651;
 Best Local Similarity 49.6%; Pred. No. 8.8e-89;
 Matches 331; Conservative 114; Mismatches 196; Indels 27; Gaps 8;

```

QY 1 MDHKLARILLDIGEGENIVAAACATRLRLVLTQTDVDRQSLDDDDPKGFEETGGM 60
DB 1 MDHKLARILLDIGEGENIVAAACATRLRLVLTQTDVDRQSLDDDDPKGFEETGGM 60
QY 61 FOIIVGPDVHVFEKLDATSKDIASVTEQKDVANANMFSRAVYLAIDFVPLIPI 118
DB 61 FOIIVGPDVHVFEKLDATSKDIASVTEQKDVANANMFSRAVYLAIDFVPLIPI 118
QY 119 PIVGGGLMAINNVLAQDLFGPQSLVEMFPQISGVAEMINMASAPFAFLPVLVGFTATK 178
DB 119 PIVGGGLMAINNVLAQDLFGPQSLVEMFPQISGVAEMINMASAPFAFLPVLVGFTATK 178

```

```

Db      119 PALVAGGLMALNNVLTAEHLFMAKSVSEVYFGLKGIAMINAMASAPFTPIILLGFGSA 178
Qy      179 TRRFGNERTLACIGAGMAYFPTLVNGYDVAATMTGEMMMSLFGLDVAQAGYGVLPV 238
Db      179 TRRFGNERTLACIGAGMAYFPTLVNGYDVAATMTGEMMMSLFGLDVAQAGYGVLPV 238
Qy      239 LVVSWILATIEKFLHRLMGTADFLITPVLTLLTGFLTFIAGPMRWGDLAAHGLQG 298
Db      239 LGVATILATLEKFFHHIKGADFPTTPEFAIVITGFLFTTVGPVLKRVSDALNGLVG 298
Qy      299 LVDFGPGVGGLLFGLVSPVITVTGLHQSFPIELFLF---NQGSGFIFATASMANIAG 354
Db      299 LVNSGMIGMGI FGLLYSAIVITGLHQTFPALETQLANVAKTGSGFIFPVA5MANIGG 358
Qy      355 AACLAVFLAKSEKLGAGASGVAVGITEPAIFGVNLRKMPYIIGTALGALIT 414
Db      359 AATLAIFFATKSKOKKALTSAGVSALGITBPAIFGVNLKMKFPFVPAALIASIASAPL 418
Qy      415 ALFDIKAVNLAAGAFLGVVSIAPDMWMLVCAVVTFFVIAFGAALAYGLVRRNGSIDP 474
Db      419 GLFPHLVAMGPASVYIGFISISKSIAPAMLSAVISFVAFTPTFIY---AKRTLGDGR 474
Qy      475 DATAAPVPAITKAEAPAEPSNDSTIIQAPLTGEAIALSSVSDAMPASGKLGSGVALV 534
Db      475 DVKSPAPITSTV-----INVDEIISAVTGSSESLKQVNDQVFAELMGKGAALV 525
Qy      535 PIRGQLVSPVSGKIYVAPPSGAFAVRTAEEDGSNDVILMHITGPTVNLNGHFNPLKQ 594
Db      536 PSSDVVAPADAVITVYDSSHAYGKTTA---GAEIIHIGLPTVNLNGHFTTNQK 581
Qy      595 GEVAKAGELLCEFDIDAKAGYEVTTPIVSNYKKTGPVNTYGGELBAGANL--LVNA 652
Db      582 GDTVAGDGLTGFTDIALAKAANYDPTVNLVITNTANYANVERLKVTNVAGSGLVALTAP 641
Qy      653 KKEAVPAT 660
Db      642 AASSVAAT 649

RESULT 3
Q88ZV9 PRELIMINARY; PRT; 651 AA.
AC Q88ZV9;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DN Sucrose PTS, EIIBCA (EC 2.7.1.69).
GN Name=pts1BCA; OrderedlocusNames=1p_0185;
OS Lactobacillus plantarum.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
NC NCB1_taxid=1590;
RN NC_011111.1
RX SEQUENCE FROM N.A.
RC STRAIN=NCIMB 8826 / WCFS1;
MEDLINE=22480296; PubMed=12566566; DOI=10.1073/pnas.0337704100;
RA Kleerebezen M., Boekhorst J., Van Katerburg R., Molenaar D.,
  Kuisberg O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
  Raaijmakers M.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
  Raaijmakers M.E.J., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,
  De Vos W.M., Sleszen R.J.;
RA "Complete genome sequence of Lactobacillus plantarum WCFS1.";
RT Proc. Natl. Acad. Sci. U.S.A. 100:11990-11995(2003).
RL HSBP; AL935252; CAD62855.1; -.
DR HSBP; AL935252; 1GPR.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; F:protein-N(P)-phosphohistidine-sugar phospho. . .; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009401; F:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011055; Dup_hybrid_motif.

```

```

DR InterPro; IPR001966; Ptrane EIIB.
DR InterPro; IPR003352; Ptrane EIIC.
DR InterPro; IPR001327; PTS EIIA.
DR InterPro; IPR010973; PTS IIB_sucr.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF00367; PTS_EIIB; 1.
DR Pfam; PF02378; PTS_EIIC; 1.
DR ProDom; PD001476; Ptrane EIIB; 1.
DR ProDom; PD002243; PTS EIIA; 1.
DR TIGRFAMs; TIGR00830; PTBA; 1.
DR TIGRFAMs; TIGR01996; PTS-II-BC-sucr; 1.
DR PROSITE; PS00371; PTS_EIIA_1; 1.
DR PROSITE; PS01035; PTS_EIIB_Cys; 1.
DR Complete proteome; Transferase.
SQ SEQUENCE 651 AA; 68514 MW; CFB535EDDB73DD6 CRC64;

Query Match      48.1%; Score 1607.5; DB 2; Length 651;
Best Local Similarity 49.3%; Pred. No. 1.2e-88;
Matches 332; Conservative 114; Mismatches 189; Indels 39; Gaps 9;

Qy      1 MDHKLQRIKLDIGEDENYVAAACATRLALVKTQDVDRQSLDDDDPDLKGTETGGM 60
Db      1 MNHQLQRIKLDIGEDENYVAAACATRLALVKTQDVDRQSLDDDDPDLKGTETGGM 59
Qy      61 FOIIVPGDVDFVFE-----LDDATSKDIAVSTEQKDVVA--NNAMWSBAVKVLAD 112
Db      60 YOIIVPGDVDFVFE-----LDDATSKDIAVSTEQKDVVA--NNAMWSBAVKVLAD 112
Qy      113 IFVPLIPLVGGGLLMAINNVLAQDLFPOSIVEMFPOISGVAMINLMASAPAPLIPV 172
Db      113 IFIPIPLVAVGGGLLMAINNVLTAEHLFMAKSVSEVYFGLKGIAMINAMASAPFTPIPI 172
Qy      173 LVGFATATGFGNERTLACIGAGMAYFPTLVNGYDVAATMTGEMMMSLFGLDVAQAGY 232
Db      173 LVGFATATGFGNERTLACIGAGMAYFPTLVNGYDVAATMTGEMMMSLFGLDVAQAGY 232
Qy      233 GTVLPVLVSWILATIEKFLHRLMGTADFLITPVLTLLTGFLTFIAGPMRWGDL 292
Db      233 GTVLPVLVSWILATIEKFLHRLMGTADFLITPVLTLLTGFLTFIAGPMRWGDL 292
Qy      292 GQVLPVLVSWILATIEKFLHRLMGTADFLITPVLTLLTGFLTFIAGPMRWGDL 292
Db      292 GQVLPVLVSWILATIEKFLHRLMGTADFLITPVLTLLTGFLTFIAGPMRWGDL 292
Qy      293 AHGLOGLVDFGPGVGGLLFGLVSPVITVTGLHQSFPIELFLF---NQGSGFIFATAS 348
Db      293 AHGLOGLVDFGPGVGGLLFGLVSPVITVTGLHQSFPIELFLF---NQGSGFIFATAS 348
Qy      293 TNGLVGLVNTTGMIGMGI FGLLYSAIVITGLHQTFPALETQLANVAKTGSGFIFPVA 352
Db      293 TNGLVGLVNTTGMIGMGI FGLLYSAIVITGLHQTFPALETQLANVAKTGSGFIFPVA 352
Qy      349 ANIAGAACLAVFLAKSEKLGAGASGVAVGITEPAIFGVNLRKMPYIIGTALGALIT 408
Db      349 ANIAGAACLAVFLAKSEKLGAGASGVAVGITEPAIFGVNLRKMPYIIGTALGALIT 408
Qy      353 ANIAGAACLAVFLAKSEKLGAGASGVAVGITEPAIFGVNLRKMPYIIGTALGALIT 412
Db      353 ANIAGAACLAVFLAKSEKLGAGASGVAVGITEPAIFGVNLRKMPYIIGTALGALIT 412
Qy      409 ICGALIALFDIKAVNLAAGAFLGVVSIAPDMWMLVCAVVTFFVIAFGAALAYGLVLR 468
Db      409 ICGALIALFDIKAVNLAAGAFLGVVSIAPDMWMLVCAVVTFFVIAFGAALAYGLVLR 468
Qy      469 NGSIDPDAITAPVPAITKAEAPAEPSNDSTIIQAPLTGEAIALSSVSDAMPASGKLG 528
Db      469 NGSIDPDAITAPVPAITKAEAPAEPSNDSTIIQAPLTGEAIALSSVSDAMPASGKLG 528
Qy      469 TLGDRODVKSPAPITSTV-----INVDEIISAVTGSSESLKQVNDQVFAELMG 519
Db      469 TLGDRODVKSPAPITSTV-----INVDEIISAVTGSSESLKQVNDQVFAELMG 519
Qy      529 SGVALVPTKQGVSPVSGKIYVAPPSGAFAVRTAEEDGSNDVILMHITGPTVNLNGH 588
Db      529 SGVALVPTKQGVSPVSGKIYVAPPSGAFAVRTAEEDGSNDVILMHITGPTVNLNGH 588
Qy      520 KGAALVFPADVAVITVYDSSHAYGKTTA---GAEIIHIGLPTVNLNGHFTTNQK 575
Db      520 KGAALVFPADVAVITVYDSSHAYGKTTA---GAEIIHIGLPTVNLNGHFTTNQK 575
Qy      589 NPLKKGGEVAKAGELLCEFDIDAKAGYEVTTPIVSNYKKTGPVNTYGGELBAGANL 648
Db      589 NPLKKGGEVAKAGELLCEFDIDAKAGYEVTTPIVSNYKKTGPVNTYGGELBAGANL 648
Qy      576 TTNVAGKGTIVAGDGLTGFTDIALAKAANYDPTVNLVITNTANYANVERLKVTNVAG 635
Db      576 TTNVAGKGTIVAGDGLTGFTDIALAKAANYDPTVNLVITNTANYANVERLKVTNVAG 635
Qy      649 LNVAKKEA--VPAT 660
Db      636 VALTEPTASSVAAT 649

RESULT 4
Q9S6S6 PRELIMINARY; PRT; 650 AA.
AC Q9S6S6;

```

DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
 DE Enzyme II sucrose protein (EC 2.7.1.69).
 GN Name=sacB;
 OS Lactococcus lactis.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1358;
 RN
 RN SEQUENCE FROM N.A.
 RX MEDLINE=99173919; PubMed=10074089;
 RT Lueink E.J., Marugg J.D., Kuipers O.P., De Vos W.M.;
 RT "Characterization of the divergent sacB and sacA operons, involved
 RT in sucrose utilization by Lactococcus lactis.";
 RL J. Bacteriol. 181:1924-1926(1999).
 DR EMBL; 297015; CAB09690.1; -.
 DR HSSP; P45618; 2GPR.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008982; P:protein-N(P)-phosphohistidine-sugar phospho. . .; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR011055; Dup hybrid motif.
 DR InterPro; IPR003352; Ptrans_EIIB.
 DR InterPro; IPR003352; Ptrans_EIIB.
 DR InterPro; IPR001127; Pts_EIIA.
 DR InterPro; IPR010973; Pts_II_BC_sucr.
 DR Pfam; PF00358; Pts_EIIA_1; 1.
 DR Pfam; PF00367; Pts_EIIB_1.
 DR Pfam; PF02378; Pts_EIIC_1.
 DR ProDom; PD001476; Ptrans_EIIB; 1.
 DR ProDom; PD002243; Pts_EIIA; 1.
 DR TIGRFAMs; TIGR00830; PTA; 1.
 DR TIGRFAMs; TIGR01996; Pts-II-BC-sucr; 1.
 DR PROSITE; PS00371; Pts_EIIA_1; 1.
 DR PROSITE; PS01035; Pts_EIIB_Cys; 1.
 KM Transferase.
 SQ SEQUENCE 650 AA; 69636 MW; 250E26F7664D204 CRC64;
 Query Match 45.1%; Score 1508; DB 2; Length 650;
 Best Local Similarity 46.9%; Pired. No. 1.1e-82;
 Matches 313; Conservative 123; Mismatches 192; Indels 40; Gaps 10;
 QY 1 MDKDLAQRIRLDIGEDNIVAAHCATRLRLVLTQKDVDRQSDDDPDLKGTETGM 60
 DB 1 NMHKQVAERILNAV-GRDNIQGARHCATRLRLVLTQKDVDRQSDDDPDLKGTETGAQG 59
 QY 61 PIIIVGGGVHDVPEKELDQATSKDIIVSTEQKDVAN--NANWFSBAKVLAIDFVPLI 118
 DB 60 YIIIVGGGVNTVVEFTYLTGISB-STADLKEIAGSKKQPMALVKLSDIFVPLI 118
 QY 119 PIVVGGGLIMAINNVLVADLDFPQSLVEMFPQISGVAMINIMASAPFAFLVVGFTA 178
 DB 119 PALVAGGLIMALNNLTAHFLATKSIVEMFPMWGFADIVTMSAAPTETPIIIGISA 178
 QY 179 TKRFGNEFLGAGIGAMVFTLVNGYDVAATMTGKEMPSLFGLDVAQAGYQGVLPV 238
 DB 179 TKRFGGNPTLGAIVGQIMVMPGLINGYVAEAISSHMTYWDIFGPKVAQAGYQGVLPV 238
 QY 229 LVVSMILATIEKFLHKRLMGTDPLITPLTLLGLFLFIAGPMRVGDLAHGQ 298
 DB 229 IGVAVILAKLERFFHKYLNDAIDFTFTPLSLVITGFLFTLVGPKLRVNSGLTDLVG 298
 QY 239 IYDFGGVGGFLFGLVYSPVITVGLHQSPPELELEFNO-----GGSFIFATASMANIA 352
 DB 239 LVNTIGALGMVFGGYSAIVTGLHQSPPELELEFNO-----GGSFIFATASMANIA 358
 QY 333 QGAACLAIVFLAKSEKTLKLAGAGVSAVLGITTEPAIFGVNRLRMPPFTIGTALGCA 412
 DB 339 QGATFAIFVTKIKTKTALAAPAGVASALIGITEPALFGINKTKYFPFIALGASAGSL 418
 QY 413 LIALPDIKVALGAAGFLGVSGIDAPDMVFLVCAVTFVIAFGAIALVGLVVRNSGI 472

DB 419 FMGLFHLVLAISLGSAGLIGFISIAKGYVLMISFISFLAFLVVTSTYGRMAK--SI 476
 QY 473 DDDATAAPVPAATTKAENAPAEFSNDSTIIQ-----APLTGEALATSSVDAMFAS 524
 DB 477 TRE-----KNKQNAITTYQPEKVIIDPVKSGELAPINGFVPLSDVSDPVFSK 525
 QY 525 KGLSGVAIVPTKQQLVSPVSGKIVVAPPSGHAARVTRKADSGSNVIDLMHIGFDYMLN 584
 DB 526 EIMGKGIATKPSGSELPSPADGEIITIAVETHAAGIKTK---NGGEVLHIGIDTVSMN 581
 QY 585 GTHFNPILKQGDDEVYAGELCEFDIDAIIKAGYVTTPIVSN--YKKTGPVNTYGI 641
 DB 582 GNGFQIVKVKQAKKADLLSPDKETIKSGLDPTVIIVITNASTYVEIIPLEN--VD 639
 QY 642 IEAGANML 649
 DB 640 IKVGEKIL 647
 RESULT 5
 ID PISA STRMU STANDARD; PRT; 664 AA.
 AC P12655;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE PTS system, sucrose-specific IIAB component (EIIABC-Scr) (sucrose-
 DE permease IIAB component) (Phosphotransferase enzyme II, ABC
 DE component) (EC 2.7.1.69) (EII-Scr).
 GN Name=scrA; OrderedLocustNames=SMU.1841;
 OS Streptococcus mutans
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1309;
 RN
 RN SEQUENCE FROM N.A.
 RC STRAIN=GS-5;
 RX MEDLINE=89123027; PubMed=2536656;
 RA Sato Y., Poy F., Jacobson G.R., Kuramitsu H.K.;
 RT "Characterization and sequence analysis of the scrA gene encoding
 RT enzyme IICer of the Streptococcus mutans phosphoenolpyruvate-dependent
 RT sucrose phosphotransferase system.";
 RL J. Bacteriol. 171:263-271(1989).
 RN
 RN SEQUENCE FROM N.A.
 RP STRAIN=UA159 / ATCC 700610 / Serotype C;
 RC MEDLINE=22295063; PubMed=12397186; DOI=10.1073/pnas.172501299;
 RX Ajdic D.J., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
 RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H.G., Lin S.P.,
 RA Qian Y., Li S., Zhu H., Najjar F.Z., Lai H., White J., Roe B.A.,
 RA Ferretti J.J.;
 RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
 RT pathogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
 RN
 RN SEQUENCE OF 639-664 FROM N.A.
 RC STRAIN=GS-5;
 RX MEDLINE=93329360; PubMed=8336109;
 RA Sato Y., Yamamoto Y., Kizaki H., Kuramitsu H.K.;
 RT "Isolation, characterization and sequence analysis of the scrA gene
 RT encoding fructokinase of Streptococcus mutans.";
 RL J. Gen. Microbiol. 139:921-927(1993).
 CC
 CC -1- FUNCTION: This is a component of the phosphoenolpyruvate-dependent
 CC sugar phosphotransferase system (PTS), a major carbohydrate active
 CC and the transmembrane channel; the IIA domain contains the primary
 CC phosphorylation site (the donor is phospho-HPr); IIA transfers its
 CC phosphoryl group to the IIB domain which finally transfers it to
 CC the sugar.
 CC
 CC -1- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
 CC histidine + sugar phosphate.
 CC
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.


```

CC -1- SIMILARITY: Contains 1 PTS_EIIB domain.
CC -1- SIMILARITY: Contains 1 PTS_EIIB domain.
CC -1- SIMILARITY: Contains 1 PTS_EIIB domain.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@sb-sib.ch).
CC -----
CC EMBL: M22711; AAA26971.1; -.
CC EMBL: AE015011; AAN59464.1; -.
CC EMBL: D13175; BAA02466.1; -.
CC FIR; B32243; B32243.
CC HSSP; P08837; 1GGR.
CC InterPro; IPR011055; Dup_hybrid_motif.
CC InterPro; IPR001996; Ptrans_EIIB.
CC InterPro; IPR003352; Ptrans_EIIC.
CC InterPro; IPR001127; PTS_EI1A.
CC InterPro; IPR010973; PTS-II-BC_sucr.
CC Pfam; PF00358; PTS_EI1A_1; 1.
CC Pfam; PF02378; PTS_EIIB; 1.
CC ProDom; PD001476; Ptrans_EIIB; 1.
CC ProDom; PD002243; PTS_EI1A; 1.
CC TIGRfam; TIGR00826; E1B_glc; 1.
CC TIGRfam; TIGR00830; PTBA; 1.
CC TIGRfam; TIGR01995; PTS-II-ABC-beta; 1.
CC TIGRfam; TIGR01996; PTS-II-BC-sucr; 1.
CC TIGRfam; TIGR01992; PTS-II-BC-tre; 1.
CC PROSITE; PS00371; PTS_EI1A_1; 1.
CC PROSITE; PS01035; PTS_EIIB_Cys; 1.
CC Complete proteome; Phosphorylation; Phosphotransferase system;
CC Sugar transport; Transferase; Transmembrane.
KW DOMAIN 1 40 EIIB.
FT DOMAIN 111 420 EIIC.
FT MOD_RES 533 664 EI1A.
FT MOD_RES 25 25 Phosphocysteine (By similarity).
FT MOD_RES 331 331 Phosphohistidine (By similarity).
FT MOD_RES 585 585 Phosphohistidine (By similarity).
SQ SEQUENCE 664 AA; 69988 MW; 8095633281A9A1 CRC64;

Query Match 44.2%; Score 1478.5; DB 1; Length 664;
Best Local Similarity 46.0%; Pred. No. 7e-81;
Matches 311; Conservative 127; Mismatches 203; Indels 35; Gaps 13;

QY 1 MOHKDLAQRILRDIGEDNIVAAAHGATRLRLVLDKTDVDRQSLDDDDPKGTFETGGM 60
DB 1 MYSKVASVITAV-GKDNLVAAAHGATRLRLVLDKDSKVDQALDKADVKGTFTDQ 59
QY 61 FOIIVGGDVHDVFKELDDATSKDIADVSTOLKDVANNA--NMFSAVKVLADIFVPLI 118
DB 60 YQVITIGPGVNFVYDEIIRKQGL-TEVSTDDLKRIASGKKRPVIALKLSIDIVPII 118
QY 119 PLVVGGLMALNNVAVADLFGPQSLVEMFQISVAENIMASAPAPFVLVGFPA 178
DB 119 PALVVGGLMALNNFLTSGLFETSLVQOFPRIKSSDMIDLMSAAPWFPPIIVGIS 178
QY 179 TRRFGNEFLAGISGMWVFTLVNGVDVA-----TWTAGEMP-WMSLFGLDVAQAY 231
DB 179 ARRFANQFLGASIGIMVAFGANIIGLAANAIPISKATTIGATGFMNIFGLHTQASY 238
QY 232 OCTVLVPLVSVILATIEFKRLKRLMGTADEITPVLTLTLGFLTFIAGPAMRWGDL 291
DB 232 TVQVLPVLAVMWLLSTIEKFFPKRLPSAVDFTFPLSLVITIGFLFIYIIGPMKEVSD 298
QY 299 LAHGQGLVDFGPGVGGLLFGLVSPITVTLGHQSPPIELEL---FNGG---GSTIAT 345
DB 299 LINGVIMVLDTTGFGMGVFGALYSPVMTGTHQSPALIEIDLISAFONGTGHGFIY 358
QY 346 ASMAIAGCAACLAFLAKSEKLGKLAGASGVSAVLGITBPALFGVNLRLMPFIYIG 405

```

```

DB 359 ASMAIAGCAACLAFLAKSEKLGKLAGASGVSAVLGITBPALFGVNLRLMPFIYIG 418
QY 406 TTAIGGALLALPDIAKVALGAAGFLGVSIDPDMWVFLCAVVTFFVIAFGAIAVGLY 465
DB 419 GSASMAAIIAGLQVAVVAGSAGFLGSLIKASIPFYVCELISPAIAFATYVG---- 475
QY 466 VRRNGSIDDAAPAPPACTTAELAEAPAEFNSDT-----IIQPLTGHAIALSVS 518
DB 476 --KTVAVDFAAEAEEAVEAEIEVQ-EIPEEAASANKAQVTDEVLAAPLAGAEVELTSV 532
QY 519 DAMFASGRKSGVAIVPTKGLVSPVSGKI VVAPSPGHAFAVTRAEAGSDNVLIMHGE 578
DB 533 DVFSEAMGKGIATKPSNTYVAPVDGTVOJAFDGHAYGI--KSDNGA--ELIHIGI 588
QY 579 DTVNLNGTHFNPLKQGEVKAAGELLCEFDIDAIKAAGYVTTPIVNSYKKTGPVNTY- 637
DB 589 DTVNSMGKGFEGQVQADQIKKGDVLTGTFDSKIAEAGLDNTMTFIVTADYASVETLA 648
QY 638 GLGETEAGANLNVAK 653
DB 649 SSGTVAVGDLSLEVRK 664

RESULT 6
ID Q7WMP7 PRELIMINARY; PRT; 649 AA.
AC Q7WMP7;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DE 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Sucrose PTS transporter (EC 2.7.1.69).
GN Name=sctra;
OS Lactobacillus acidophilus.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxId=1579;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCFW;
RX PubMed=12847288; DOI=10.1073/pnas.1332765100;
RA Barrangou R., Altermann E., Hutkins R., Cano R., Klaenhammer T.R.;
RT "Functional and comparative genomic analyses of an operon involved in
RT fructooligosaccharide utilization by Lactobacillus acidophilus."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8957-8962(2003).
DR EMBL; AY177419; AA038866.1; -.
DR HSSP; P20166; 1AX3.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; P:protein-N(PI)-phosphohistidine-sugar phospho. . .; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0016740; P:transferase activity; IEA.
DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. . .; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011055; Dup_hybrid_motif.
DR InterPro; IPR001996; Ptrans_EIIB.
DR InterPro; IPR001127; PTS_EI1A.
DR InterPro; IPR010973; PTS-II-BC_sucr.
DR Pfam; PF00358; PTS_EI1A_1; 1.
DR Pfam; PF00367; PTS_EIIB; 1.
DR ProDom; PD001476; Ptrans_EIIB; 1.
DR ProDom; PD002243; PTS_EI1A; 1.
DR TIGRfam; TIGR00830; PTBA; 1.
DR TIGRfam; TIGR01996; PTS-II-BC-sucr; 1.
DR PROSITE; PS00371; PTS_EI1A_1; 1.
DR PROSITE; PS01035; PTS_EIIB_Cys; 1.
KW Transferase.
SQ SEQUENCE 649 AA; 69446 MW; B4377CA24D00507B CRC64;

Query Match 43.6%; Score 1457; DB 2; Length 649;
Best Local Similarity 45.0%; Pred. No. 1.3e-79;
Matches 290; Conservative 122; Mismatches 191; Indels 42; Gaps 10;

QY 1 MOHKDLAQRILRDIGEDNIVAAAHGATRLRLVLDKTDVDRQSLDDDDPKGTFETGGM 60

```

D	R	GO: 0009480; P: phosphoenolpyruvate-dependent sugar phospho. . ; IEA.
D	R	GO: 0006810; P: transposon; IEA.
D	R	InterPro: IPR011055; Dup hybrid_motif.
D	R	InterPro: IPR015835; E1IB_glc.
D	R	InterPro: IPR001996; Petam_E1IB.
D	R	InterPro: IPR003352; Petam_E1IC.
D	R	InterPro: IPR001127; Pts_E1A.
D	R	InterPro: IPR010973; Pts II BC_sucr.
D	R	Pfam: PF00358; Pts_E1A_1; 1.
D	R	Pfam: PF00367; Pts_E1IB_1.
D	R	Pfam: PF02378; Pts_E1IC_1.
D	R	Prodrom: PD001476; Petam_E1IB; 1.
D	R	Prodrom: PD002243; Pts_E1IA; 1.
D	R	TIGRFAMs: TIGR00826; E1B_glc; 1.
D	R	TIGRFAMs: TIGR00830; PTBA; 1.
D	R	TIGRFAMs: TIGR01996; Pts-II-BC-sucr; 1.
D	R	PROSITE: PS00371; Pts_E1A_1; 1.
D	R	PROSITE: PS01035; Pts_E1IB_Cys; 1.
K	M	Complete proteome; Pyruvate; Transferrase.
S	Q	SEQUENCE 647 AA; 69202 MW; 95FA7AC53CB9541 CRC64;
		Query Match 43.3%; Score 1446.5; DB 2; Length 647;
		Best Local Similarity 45.0%; Pred. No. 5.8e-79;
		Matches 300; Conservative 125; Mismatches 209; Indels 33; Gaps 9
Qy		1 MDKDLAQRIIRDIGGEDNIVAAAHGATRLRLVTKDTKDVRDQSIDDPPDKGTETGGM 60
Dd		1 MDHKKVAAVEAV-GRDNLVAAAHGATRLRLVLDDSKVMKKMLSDSPDIGTIRKNGQ 59
Qy		FOIITGPEDVDHVFELDATSKOITAVSTEOIKDVANNA--NWFSRAVKVIADIFVPIL 118
Dd		FOIITIGPDVANNYDELIKMTGLS-ELSTIDLKRYAKOQDKENPMAPAFIKLSDIFVPII 118
Qy		PLTVGGGLIMLAINNVLVAODLFEGPOSIVEMFPPOISGVAMINLMASAPPALFVLVGFLTA 178
Dd		PALVAGGLMALNNLFTSPGLFGAKSVQMAPNVNGLSEIIQVMSAAPFIEMPIIVGMSA 178
Qy		TKRRCGNRPFDAGCGMAWPFPLVUNGVDAVTMTGEMPMWSLFGLDVAQGYOGVTLRV 238
Dd		AARFRANPFLGVTIGMINTTPAL-----GGAQRNDIFGLHSQNTNYQOVITIV 227
Qy		LVSWSIILATIEFKLRIMGTADFLLTFVLLLTGPTFLFIAGMRWGDLLHGLQG 298
Dd		LVAWVLALFLERFRPKWLPSAVDFEFTPLDSIMTGFLTFTIIGPVFKVSADAITAIWM 287
Qy		LYDEGGPVGILLFGLVSPYIVITGHOSPPRIELEFNQ-----GGSRTFAPTASANTA 352
Dd		LYNTTGAAGMGIFGLSYSAIVTTGHOSPRAVETOLLAEVARGSGDPIFYTACMANVA 347
Qy		QGAACIAFVFLAKSEKLKGLGASGSVASVLGITTEPAIFGVNLRLRWPFYIGITLAIGA 412
Dd		GGAATFAIYFLTKNKXVGLASSGSVALLGITTEPALFGVNLYKYFFPCALIGAVGAA 407
Qy		LIALFDIAVALGAAGFLGVSVSIDAPDMVMFLVCVVTFVIAFGAAILAGLY-----LYR 467
Dd		FAGLMHWVAALGASGFGLFSMWKSJIPMALSVISIEIVAFGLTFYVKGKHFEKDEIVE 467
Qy		RNGSIDPDATAVPVAGTTKEABAPAEBSNSDSTIOAPLTGEALIASVSDAMFASGKL 527
Dd		ESGTYE---SADQVAOQEKAQQIIEKDEKLDEVIYAAPVSGASKASIKMDNPVISTEAM 524
Qy		SGSVAIVPTKQOLVSPVSGKIYVAFPSSGAFAVRTKAEDGSNVDILMHIGPTVLINGTH 587
Dd		GKGAAMVPESEGIYSPVTGEITVAAYETKAYGI--KSDEGA--EVLIIHGIDTVMNKKGG 580
Qy		FNPLKQGDENVAAGBLLCEFDIDALKAGYEVTTPILVSNYKKTGTPVNVYGIGEIEAGAN 647
Dd		FTTDVQOQHVERKSELGTVIDDAVKKAGYDTTMVITNTTSYANQRIIDOVERKKGDD 640
Qy		LUNVAKK 654
Dd		LIAVTKR 647

[illegible]

```

RESULT 9
G93ML1
ID Q93ML1 PRELIMINARY; PRT; 534 AA.
AC Q93ML1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Sucrose-specific enzyme II of the PTS (Fragment).
GN Name=scria;
OS Lactobacillus sakei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1599;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21822499; PubMed=11832506;
RA Dutez A.M., Chailou S., Hissler V., Stentz R.,
RA Champomier-Verges M.C., Albert C.A., Zagrec M.;
RT "Physical and genetic map of the Lactobacillus sakei 23k chromosome.";
RN Microbiology 148:421-431(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Dutez A.-M., Chailou S., Zagrec M.;
RL Submitted (JULY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF401046; AAK92528.1; -.
DR HSSP; P45618; 2GPR.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008982; C:protein-N(P)-phosphohistidine-sugar phospho. . ; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0009401; F:phosphoenolpyruvate-dependent sugar phospho. . ; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR011055; DUF_hybrid_motif.
DR InterPro; IPR003552; Pfam_EIIC.
DR InterPro; IPR01127; PTS_EIIA.
DR Pfam; PF00358; PTS_EIIA_1; 1.
DR Pfam; PF02378; PTS_EIIC_1.
DR ProDom; PD002243; PTS_EIIA_1.
DR TIGRFAMs; TIGR00630; PTBA; 1.
DR PROSITE; PS00371; PTS_EIIA_1; 1.
FT NON TER 1

```

SQ SEQUENCE 534 AA; 56230 MW; 01804F9DE70C0089 CRC64;
 Query Match 37.8%; Score 1264; DB 2; Length 534;
 Best Local Similarity 47.6%; Pred. No. 4,6e-68;
 Matches 265; Conservative 99; Mismatches 155; Indels 38; Gaps 10;

114 FVPLPIIVGGGLMANNVVAQDLFGPQSLVEMFQISGVAMINIMASAPAFPLVL 173
 1 FVPLPIIVGGGLMANNVVAQDLFGPQSLVEMFQISGVAMINIMASAPAFPLVL 60
 174 VGFITKTRRGNEFGAGIGAMVPTLVNGVDVAAITAGMPMWSLFGLDVAQAGOG 233
 61 IAFSRTKRRGNPYGAAAGMWLVMPNIVNGVABSIATGHTYWHVFGALIAQAGOG 120
 234 TVLPVLVSMILATIEKFLHKLMTADPLTPVTLTLTGFLTFIAIGPAMRWGDLLA 293
 121 QVPIPIGAFILANIEKFPKHLNDVADPTFPMISIIITGTLTIVGPAIRIVSNVT 180
 294 HGLQGLYDFGSPVGGLLFGLVSPVITVTLGHOSPPIELFLF---NOGGSFIPTASMA 349
 181 DSLVMAVQTLGAVGNGIFGLGSAIVLTGLHOSFPAIETTLADIATKGGSFIFVAMA 240
 350 NIAQGAACLAVERFLAKSEKLKGLAGSGVSAVLGITEPAIFGVNLRMPFYIGITLAI 409
 241 NIAQGAATPAVFVTVKXKQKSLTTSAGISANGLTEPFLFVNLKLPFPFGLIAGSI 300
 410 GGALIALPDIKAVALGAGFLGVSDIDAPDMVFLVCAVVFVIAFGAIAVGLYVRN 469
 301 SSFIIIGLHLVLSVSMGPACIGFIALAPKSTPSFMMAIISFVIAFVGTLYGKKAMK-- 358
 470 GSIDPATAAPVPACTTKEA--EAPA-----EFSNDSTIIQAPLTGELIALSSVDMF 522
 359 -----TTEEEIINEAPATPEVERVLDK--ISAPVGRVLDIASVPDPVF 402
 522 ASGKSGVAIVPTKQOLSPVSGKIVAFPSGHAFAVTRKADSGSNVDILHIGFETVN 582
 403 ABEAMKGLAIMPDSQDVLAIPVGTITIAANTGHAIGI--KSDDA--EVLHIGIDITN 458
 583 LNGTHENPLKKQDEYKAGELLCEPDIDAIKAAIGEVTTPIVSN--YKKTGPVNTYGL 639
 459 LNGIGREKIVQOGQHVSEBGLLGHFIDIKIKOAGLPLMTITVMTAGVADPDLITVDK 518
 640 GEIEGAMNLLN-AKKE 655
 519 AAMQ-GEETIQLAKKD 534

RESULT 10
 PTSB_STAXY STANDARD; PRT; 480 AA.
 ID PTSB_STAXY STANDARD; PRT; 480 AA.
 AC P51184;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE PTS system, sucrose-specific IIBC component (EIIBC-Scr) (Sucrose-
 permease IIBC component) (Phosphotransferase enzyme II, BC component)
 DE (EC 2.7.1.69) (EII-Scr).
 GN Name=scrA;
 OS Staphylococcus xylosum.
 OC Bacteria; Firmicutes; Bacilliales; Staphylococcus.
 OX NCBI_TaxId=1288;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 20267 / Isolate C2A;
 RA MEDLINE=94049686; PubMed=8232209;
 RX Wagner E., Goetz F., Bruckner R.;
 RT Cloning and characterization of the scrA gene encoding the sucrose-
 specific Enzyme II of the phosphotransferase system from
 Staphylococcus xylosum.";
 RT Mol. Gen. Genet. 241:33-41(1993).
 CC -1- FUNCTION: This is a component of the phosphoenolpyruvate-dependent
 sugar phosphotransferase system (PTS), a major carbohydrate active
 transport system. The IICD domains contain the sugar binding site

CC and the transmembrane channel; the IIA domain contains the primary
 CC phosphorylation site (the donor is phospho-HPr); IIA transfers its
 CC phosphoryl group to the IIB domain which finally transfers it to
 CC the sugar.
 CC -1- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
 CC histidine + sugar phosphate.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -1- SIMILARITY: Contains 1 PTS EIIb domain.
 CC -1- SIMILARITY: Contains 1 PTS EIIc domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by, and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----

DR EMBL; X69800; CAA49461.1; -.
 DR PIR; S39978; S39978.
 DR InterPro; IPR001996; Pirans_EIIB.
 DR InterPro; IPR003352; Pirans_EIIC.
 DR InterPro; IPR010973; PTS_II_BC_sucr.
 DR Pfam; PF00367; PTS_EIIB_1.
 DR Pfam; PF02378; PTS_EIIC_1.
 DR ProDom; PD001476; Pirans_EIIB_1.
 DR TIGRFAMs; TIGR00826; EIIB_glc_1.
 DR TIGRFAMs; TIGR01995; PTS-II-ABC-beta_1.
 DR TIGRFAMs; TIGR01996; PTS-II-BC-sucr_1.
 DR TIGRFAMs; TIGR01992; PTS-IIBC-Tre_1.
 DR PROSITE; PS01035; PTS_EIIB_CYS_1.
 KM Inner membrane; Phosphorylation; Phosphotransferase system;
 KM Sugar transport; Transferase; Transmembrane.
 FT DOMAIN 1
 FT DOMAIN 2
 FT MOD RES 26
 FT MOD RES 325
 FT MOD RES 325
 SEQUENCE 480 AA; 51326 MW; AB4EID9785D84E47 CRC64;

Query Match 36.7%; Score 1227; DB 1; Length 480;
 Best Local Similarity 52.7%; Pred. No. 7e-66;
 Matches 252; Conservative 83; Mismatches 133; Indels 10; Gaps 6;

1 MDHDLAQRIRLDIGGEENIVAAACATRLVTKDTCDVDRQSLDDPDLKFTETGGM 60
 1 MMYKKSANILQALGGEENVEAMTRCATRLVTKDEGLVDEKALGDMVDVKKGTFSGCG 60
 61 FOITVPGDVDFVPEKELDDATSKDIAVSTEQDKDVANNANWFSRAVKLTADIFVPLIP 120
 61 YQVILGSGTVKVFSELEKITGKE-ASSVSEVKTQGTGNMPPFORFVMTLSIDIFVPLIP 119
 121 LVGGGLMANNVVAQDLF-GPQSLVEMFQISGVAMINIMASAPAFPLVAVGFTAT 179
 120 IVAGGLMGINNILLAPGIFPDNQSLIEVQFSGLAEMINIFANAPFTLLPILIGSAA 179
 180 KRFGENEFLGAGIGAMVPTLVNGVDVAAITAG-EMPMWSLFGLDVAQAGOGTVLPV 238
 180 KRFGENATLGAALGMIIVHPELMSAYDYPKALEKKEIPHNHNLGELINGVQGOQVLPW 239
 239 LVVSMILATIEKFLHKLMTADPLTPVTLTLTLGFLTFIAIGPAMRWGDLLAHGIG 298
 240 LVATVTLATIEKGLKAVIPVLDNLTPLLAIIISGFTFSFVGTLTTLTYWLSIDGLTW 299
 299 LVYRGPGVGLLFGIVASPIVITGHSPPLEELF-----NOGGSFIPTASMANIAQG 354
 300 IYFEGGALGILFGLVAPVITGHSFALIEVQLDSSSTGSGSIFPLAINSNIAQG 359
 355 AACLAVERFLAK-SEKLKGLAGSGVSAVLGITEPAIFGVNLRMPFYIGITLAIAGAL 413
 360 AALAAFPILIKENKTLKGVASAGSALGITEPAMFVNLRMPFGALVSGISGAY 419
 414 IALFDIKAVAGAGFLGVSDIDAP--MMVFLCAVTVFYIAGALVGLYVRN 469

Db 420 IAFKVKALATAGT:PGFISISGQNGMLHYGIAMIATVYAGVTALRYKKRN 477

RESULT 11

08CN82 PRELIMINARY; PRT; 481 AA.

AC 08CN82; 08CN82; 01-MAR-2003 (TEMBLrel. 23, Created)

DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)

DE 01-MAR-2004 (TEMBLrel. 26, Last annotation update)

OS OrderedLocNames=SE1959;

GN Staphylococcus epidermidis.

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

NCBI_TaxId=1282;

OK NCB1

RN 11

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 12228;

RX PubMed=12950922;

RA Zhang Y.-Q., Ren S.-X., Li H.-L., Wang Y.-X., Fu G., Yang J., Qiu Z.-O., Miao Y.-G., Wang W.-Y., Chen R.-S., Shen Y., Chen Z., Yuan Z.-H., Zhao G.-P., Qu D., Danchin A., Wen Y.-M.;

RT "Genome-based analysis of virulence genes in a non-biofilm-forming Staphylococcus epidermidis strain (ATCC 12228).";

RT Mol. Microbiol. 49:1577-1593(2003).

DR EMBL: AB016750; AAC05600.1; -.

DR HSPF: P05053; IIBA.

DR GO: GO:0016020; C:membrane; IEA.

DR GO: GO:0008982; F:protein-N(pi)-phosphohistidine-sugar phospho. .; IEA.

DR GO: GO:0005351; F:sugar porter activity; IEA.

DR GO: GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.

DR GO: GO:0006810; P:transport; IEA.

DR InterPro: IPR011535; E1B_glc.

DR InterPro: IPR001996; Ptrans_E1B.

DR InterPro: IPR003352; Ptrans_E1B.

DR InterPro: IPR010973; PTS_II_BC_sucr.

DR Pfam: PF00367; PTS_E1B_1.

DR Pfam: PF02378; PTS_E1IC_1.

DR ProDom: PD001476; Ptrans_E1B; 1.

DR TIGRFAMs: TIGR00826; E1B_glc; 1.

DR TIGRFAMs: TIGR01996; PTS-II-BC-sucr; 1.

DR PROSITE: PS01035; PTS_E1B_CYS; 1.

KW Complete proteome.

SO SEQUENCE 481 AA; 51609 MW; 4472BFE3B04F409 CRC64;

Query Match 35.5%; Score 1186; DB 2; Length 481;

Best Local Similarity 49.9%; Pred. No. 2.1e-63;

Matches 225; Conservative 95; Mismatches 123; Indels 18; Gaps 6;

Qy 1 MHKDLAQRIILDIGEDENIVAAACATRLRLVKTQDVROSLDDDDPKGTFTGGM 60

Db 1 MSYKQSAEBILKAIIGEBENIDMAHCAIRLRLVNDSEKVDDELTSMNVVVGKSTGQ 60

Qy 61 FQIIVPGVDVDFKELDQATSKDIAVSTEQLDKDVANNANMFSAVKVLADI FVPLPI 120

Db 61 YQIISGIVNKNVFELEKLTKE-ASTTSVKDSSKKNMFPQKVKMLSDIFVPIIA 119

Qy 121 LVGGGLMAINNVLAQDL-FGQSLVEMFPQISGAEMINIMASAPFAPLVGFTAT 179

Db 120 IYAGGLMGILNNIIFAKDLFYDOKSIIDVHGSFGSLADMINIFANAPFTLLPILIGFSA 179

Qy 180 KRFGNNEFGAGIGAMVPEPTLVNGVDVAATMTAGE-MPMMSLFGLDVAQAGYGVLPV 238

Db 180 KRFGNPFLGALGMILVHPGMSADDPKALBEGKAIPIHMDVFGAHINEVGQGVLPW 239

Qy 239 LVVSWILATIEKFLHRLMGTAFLITPVLTLLTGLFTFLAIGPMRWGDLAHLGQ 298

Db 240 LVATYILATIEKFLKRVIPVVDNLTPLLSITFTITFLVGPVTRQLGWSGLTW 299

Qy 299 LVDFGGVGGLLFGLVSPVITVTHQSPPIELF---NOGGSFIFATASMANIAG 354

Db 300 LVDFGGVGGLLFGLVSPVITVTHQSPPIELF---NOGGSFIFATASMANIAG 359

Qy 355 AACLAVEPLAK-SEKLGKLAGSASVAVGITEPAIPGNLRLRMPFYIGTAATGAL 413

Db 360 GALAAFPITIKONKKLKGVASAGISALGITPEAMPFVNLKTRYPFGIAGSGIGAA 419

Qy 414 IALFDIKAVLAGAAGFLGVASIDAP-----MWFLVCAVTFVIA 454

Db 420 ISFFKVKALATAGLPGFISINPTHAGMLHYLIGMLIAFVSVVTVLIS 470

RESULT 12

08NV35 PRELIMINARY; PRT; 480 AA.

AC 08NV35; 08NV35; 01-OCT-2002 (TEMBLrel. 22, Created)

DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)

DE PTS system sucrose-specific IIBC component.

GN Name=sucra; OrderedLocNames=NM2299;

OS Staphylococcus aureus (strain NM2).

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

NCBI_TaxId=196620;

OK NCB1

RN 11

RP SEQUENCE FROM N.A.

RC MEDLINE=22040717; PubMed=12044378; DOI=10.1016/S0140-6736(02)08713-5;

RX Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A., Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L., Yamamoto K., Hiramatsu K.;

RT "Genome and virulence determinants of high virulence community-acquired MRSA.";

RT Lancet 359:1819-1827(2002).

DR EMBL: AB004830; BAB96164.1; -.

DR GO: GO:0016020; C:membrane; IEA.

DR GO: GO:0008982; F:protein-N(pi)-phosphohistidine-sugar phospho. .; IEA.

DR GO: GO:0005351; F:sugar porter activity; IEA.

DR GO: GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.

DR GO: GO:0006810; P:transport; IEA.

DR InterPro: IPR011535; E1B_glc.

DR InterPro: IPR001996; Ptrans_E1B.

DR InterPro: IPR003352; Ptrans_E1B.

DR InterPro: IPR010973; PTS_II_BC_sucr.

DR Pfam: PF00367; PTS_E1B_1.

DR Pfam: PF02378; PTS_E1IC_1.

DR ProDom: PD001476; Ptrans_E1B; 1.

DR TIGRFAMs: TIGR00826; E1B_glc; 1.

DR TIGRFAMs: TIGR01996; PTS-II-BC-sucr; 1.

DR PROSITE: PS01035; PTS_E1B_CYS; 1.

KW Complete proteome.

SO SEQUENCE 480 AA; 51218 MW; B3EBE81C08C0DEDA CRC64;

Query Match 35.4%; Score 1183; DB 2; Length 480;

Best Local Similarity 50.4%; Pred. No. 3.1e-63;

Matches 242; Conservative 94; Mismatches 130; Indels 14; Gaps 7;

Qy 1 MHKDLAQRIILDIGEDENIVAAACATRLRLVKTQDVROSLDDDDPKGTFTGGM 60

Db 1 MSYKQSAEBILKAIIGEBENIDMAHCAIRLRLVNDSEKVDDELTSMNVVVGKSTGQ 60

Qy 61 FQIIVPGVDVDFKELDQATSKDIAVSTEQLDKDVANNANMFSAVKVLADI FVPLPI 120

Db 61 YQIISGIVNKNVFELEKLTKE-ASTTSVKASAKNMPLQGFVKMLSDIFVPIIA 119

Qy 121 LVGGGLMAINNVLAQDL-FGQSLVEMFPQISGAEMINIMASAPFAPLVGFTAT 179

Db 120 IYAGGLMGILNNIIFAKDLFFSGKSLIDVYSQFAGIAEMINIFANAPFTLLPILIGFSA 179

Qy 180 KRFGNNEFGAGIGAMVPEPTLVNGVDVAATMTAGE-MPMMSLFGLDVAQAGYGVLPV 238

Db 180 KRFGNPFLGALGMILVHPGMSADDPKALBEGKAIPIHMDVFGAHINEVGQGVLPW 239

Qy 239 LVVSWILATIEKFLHRLMGTAFLITPVLTLLTGLFTFLAIGPMRWGDLAHLGQ 298

Db 240 LVATYILATIEKFLKRVIPVVDNLTPLLSITFTITFLVGPVTRQLGWSGLTW 299

QY 299 LYDFGPGVGLFGLVYSPVITVITGLHOSFPPIELF-----NOGGSFIPATASMANIAOG 354
 Db 300 LYDFGPGVGLFGLVYSPVITVITGLHOSFPPIELF-----NOGGSFIPATASMANIAOG 359
 QY 335 AACLAFFFLAK-SEKTKGLAGASGVSAVLGITTEPAIFGVNLRMPFYIGTAAIGAL 413
 Db 360 GAATAAPFIIKONKKLKGVASAGISALLGITTEPAMGVNLRMPFYIGTAAIGAL 419
 QY 414 TALFDIKAVLGAAGFLGVSIDA--PDWVFLVCAVTVFYIAFGAIAVGLYVRNGS 471
 Db 420 IAFKVAALGTLGTLGPGFISINPVHAGMLHFVGMTISFII---ATVTLILSKKAN 475

RESULT 13

06G6U2 PRELIMINARY; PRT; 480 AA.

AC 06G6U2
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE PTS system, sucrose-specific IIBC component (EC 2.7.1.69).
 GN OrderedLocNames=SA52269;
 OS Staphylococcus aureus (strain MSSA476).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=282459;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Holden M.T.G., Fell E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
 Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
 Baason N., Bentley S.D., Chillingworth C., Chillingworth T.,
 Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
 Felwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
 James K.D., Leonard N., Line A., Mayes R., Moulé S., Mungall K.,
 Ormond D., Quail M.A., Rabinowitch B., Rutherford K.M., Sanders M.,
 Sharp S., Simmonds M., Stevens K., Whitehead S., Barrett B.G.,
 Spratt B.G., Parkhill J.;
 RT "Complete genomes of two clinical Staphylococcus aureus strains:
 evidence for the rapid evolution of virulence and drug resistance.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
 DR EMBL; BX571857; CAG4082.1; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008982; P:protein-N(P)-phosphohistidine-sugar phospho. .; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0016740; P:transferase activity; IEA.
 DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro: IPR011535; E1B_glc.
 DR InterPro: IPR001996; Ptrans_E1B.
 DR InterPro: IPR003352; Ptrans_E1B.
 DR InterPro: IPR010973; PTS_II_BC_sucr.
 DR Pfam; PF02378; PTS_E1B; 1.
 DR Pfam; PF02378; PTS_E1B; 1.
 DR ProDom; PD001476; Ptrans_E1B; 1.
 DR TIGRfam; TIGR01966; PTS_II_BC_sucr; 1.
 DR TIGRfam; TIGR01966; PTS_II_BC_sucr; 1.
 DR PROSITE; PS01035; PTS_E1B_Cys; 1.
 KW Complete proteome; Transferase.
 SQ SEQUENCE 480 AA; 51218 MW; B3EBE81C08CODEDA CRC64;

Query Match 35.4%; Score 1183; DB 2; Length 480;
 Best Local Similarity 50.4%; Pred. No. 3.1e-63;
 Matches 242; Conservative 94; Mismatches 130; Indels 14; Gaps 7;

QY 1 MDHKLADLRIDIGEDNIVAAACATRLRLVKDTKVDVRSQSLDDPDLKGTETGGM 60
 Db 1 MMYKGAHEIINALIGEENLDMAHKATRLRLVINDSELVNEALNNMVGVTSTGGQ 60
 QY 61 FOITVGPVDVHFKELDADTSKDIAVSTEQLDKDVANNANFSAVKYLAIFVPLIPI 120
 Db 61 YOIITGSGTVNKVFSELEKLTGKE-ASTTSEVAQASAKMNPLOKRVKMLSDIFVPIIPA 119

QY 121 LVGGGLMANNVIVADL-FGPOSIVMEFPQISGVAEMINIMASAPAFPLPVVGFTAT 179
 Db 120 IVAAGLGMANNITLADLFFSGSLIDVYQFGLAEMINIVFNAPFTLPIITIGSAA 179
 QY 180 KRFGNEFLAGIGAMVFPFLVNGVDVAATMTAGE-MPMMSLSEGLDVAAGYOGVLPV 238
 Db 180 KRFGNEFLAGLGMILVHSLMAYDFPKAVENGKAIPIWDVGLHINQVGGVLP 239
 QY 239 LVSMILATTEKFLHKLMTADLTPVTLTLITGLFTLTAIGANRWGDLAHLQOG 298
 Db 240 LVAAYILASIEKGRKXVIFVLDMLTLPLISFTAFLETFSFVGPIRROGLWYLSDLTW 299
 QY 299 LYDFGPGVGLFGLVYSPVITVITGLHOSFPPIELF-----NOGGSFIPATASMANIAOG 354
 Db 300 LYDFGPGVGLFGLVYSPVITVITGLHOSFPPIELF-----NOGGSFIPATASMANIAOG 359
 QY 335 AACLAFFFLAK-SEKTKGLAGASGVSAVLGITTEPAIFGVNLRMPFYIGTAAIGAL 413
 Db 360 GAATAAPFIIKONKKLKGVASAGISALLGITTEPAMGVNLRMPFYIGTAAIGAL 419
 QY 414 TALFDIKAVLGAAGFLGVSIDA--PDWVFLVCAVTVFYIAFGAIAVGLYVRNGS 471
 Db 420 IAFKVAALGTLGTLGPGFISINPVHAGMLHFVGMTISFII---ATVTLILSKKAN 475

RESULT 14

06G6S4 PRELIMINARY; PRT; 480 AA.

AC 06G6S4
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE PTS system, sucrose-specific IIBC component (EC 2.7.1.69).
 GN Name=STRA; OrderedLocNames=MSA2466;
 OS Staphylococcus aureus (strain MSA252).
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 OX NCBI_TaxID=282458;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Holden M.T.G., Fell E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
 Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,
 Baason N., Bentley S.D., Chillingworth C., Chillingworth T.,
 Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
 Felwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
 James K.D., Leonard N., Line A., Mayes R., Moulé S., Mungall K.,
 Ormond D., Quail M.A., Rabinowitch B., Rutherford K.M., Sanders M.,
 Sharp S., Simmonds M., Stevens K., Whitehead S., Barrett B.G.,
 Spratt B.G., Parkhill J.;
 RT "Complete genomes of two clinical Staphylococcus aureus strains:
 evidence for the rapid evolution of virulence and drug resistance.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
 DR EMBL; BX571856; CAG41448.1; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008982; P:protein-N(P)-phosphohistidine-sugar phospho. .; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro: IPR011535; E1B_glc.
 DR InterPro: IPR001996; Ptrans_E1B.
 DR InterPro: IPR003352; Ptrans_E1B.
 DR InterPro: IPR010973; PTS_II_BC_sucr.
 DR Pfam; PF02378; PTS_E1B; 1.
 DR Pfam; PF02378; PTS_E1B; 1.
 DR ProDom; PD001476; Ptrans_E1B; 1.
 DR TIGRfam; TIGR01966; PTS_E1B_glc; 1.
 DR TIGRfam; TIGR01966; PTS_II_BC_sucr; 1.
 DR PROSITE; PS01035; PTS_E1B_Cys; 1.
 KW Complete proteome.
 SQ SEQUENCE 480 AA; 51176 MW; EDE9B97031C26A12 CRC64;

Query Match 35.3%; Score 1179; DB 2; Length 480;
 Best Local Similarity 50.2%; Pred. No. 5.5e-63;

Matches	24;	Conservative	95;	Mismatches	130;	Indels	14;	Gaps	7
Qy	1	MDHKDLQRIIRDIGGEDNYIAAAHCATRLRLVLTQDKVDNRQSLDDDDDLKGTETGGM	60						
Db	1	MYKQASDIDLINALIGEEENLDDMAHCAITRLRLVYNDESILVNEEALANNMDVVGGTSTGQ	60						
Qy	61	FOIIVGPDGVHVFVELDDATSKDIAVSTEQLDKDVANNANMFSAVKVLIAPLPPI	120						
Db	61	YOIIIGSGTVKVFSELEKLTGKE-ASTTSERKQASAKAMFVQKPKMLSDIFPILIPA	119						
Qy	121	LVGGGLMAINNVLVAODL-FGGQSLVENEFPQISGVAEMINIMASAPPAFLVLVGTAT	179						
Db	120	IVAGGLMLGLNNLITAKDLFFSGKSLIDVYSQFAGLAEMINVFANAFTLITLIGFSNA	179						
Qy	180	KRFGNEFLGAGIGAMVPEPTLVNXYDVAATMTAGE-PMWMSFLGLDVAQNGYQSTVLPV	238						
Db	180	KRFGNPFPLGALNGILVHPSLMSAYDEPKKAVEAGALPYWDFGHLINQVYQGOQVLPW	239						
Qy	239	LVSMNLTATIEFKLHKRLMGTDPLITPVLTLTLTGFTFLAIGBMEWVGDLHAHGLQ	298						
Db	240	LVAATIIASIEKGLKRVIPVLVDNLITPLLSIFITAPLTFSPVGRITRQLDGLWLSDGILTW	299						
Qy	299	LYDEGPGVGLLFGIVSPPIVITGHOSPPRIEELF----NQGSSTFATASMANINQ	354						
Db	300	LYERFGAIGGLFGLLVAPIVITGMHSHIAVETLLADATYTGSSFLPPIATWMSNAQ	359						
Qy	355	AACLAVFPLAK-SEKTLAKIAGAGSVAVLGITPEPAIFGVNLRLMPFYIGITAGAL	413						
Db	360	GAIAAFPIIKQNKTLKGVASAAGISALLGITPEPMFGVNLKLRPIFGALVGSIGSAY	419						
Qy	414	IALPIKAVALGAAGFLGVSTDA--PDMWFLVCAVTFVYAFGAALAYGLVYVRNGS	471						
Db	420	IAFFVVKVIALGTAGLPGFISINPVAAGMLHAFVMTGISFY----ATVTLILSKRAN	475						

RESULT 15			
099R00			
ID	099R00	PRELIMINARY;	PTS; 480 AA.
AC	099R00;		
DT	01-JUN-2001 (TREMBlrel, 17, Created)		
DT	01-JUN-2001 (TREMBlrel, 17, Last sequence update)		
DT	01-MAR-2004 (TREMBlrel, 26, Last annotation update)		
DE	PTS system, sucrose-specific IICG component.		
GN	Name=sctra, orderedLocusNames=SNV2377;		
OS	Staphylococcus aureus (strain M50 / ATCC 700699).		
OC	Bacteria; Firmicutes; Bacillales; Staphylococcus.		
OX	NCBI_TaxId=158878;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	SPRAIN-Mu50 / ATCC 700699;		
RX	MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;		
RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,		
RA	Cui L., Oguchi A., Aoki K.-i., Nagai Y., Iian J.-Q., Ito T.,		
RA	Kanamori M., Matsunaru H., Maruyama A., Murakami H., Hosoyama A.,		
RA	Mitsunori-Ui Y., Takahashi N.K., Sawano T., Inoue R.-i., Kaito C.,		
RA	Selimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,		
RA	Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,		
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;		
RT	"Whole genome sequencing of methicillin-resistant Staphylococcus		
RT	aureus."		
RL	Lancef 357:1225-1240(2001).		
DR	EMBL; AP003365; BABS85539.1; -.		
DR	PIR; D90038. D90038.		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	GO; GO:0008962; F:protein-N(Pi)-phosphohistidine-sugar phospho. ; IEA.		
DR	GO; GO:0005351; F:sugar porter activity; IEA.		
DR	GO; GO:0009431; P:phosphoenolpyruvate-dependent sugar phospho. ; IEA.		
DR	GO; GO:0006810; P:transport; IEA.		
DR	InterPro; IPR011535; E1IB_glc.		
DR	InterPro; IPR001966; P1rans_E1IB.		
DR	InterPro; IPR003354; P1rans_E1IC.		
DR	InterPro; IPR010973; PTS_II_BC_sucr.		
DR	Pfam; PF00367; PTS_E1IB; 1.		

```
DR Pfam; PF02378; PTS_EIIC; 1.  
DR ProDom; PD001476; Petans_EIIB; 1.  
DR TIGRFAMs; TIGR00826; EIIB_glic; 1.  
DR TIGRFAMs; TIGR01996; PTS-II-BC-sucr; 1.  
DR PROSITE; PS01035; PTS_EIIB_CYS; 1.  
KM Complete proteome.  
SQ SEQUENCE 480 AA; 51232 MW; D88607F6E0AF2EB4 CRC64;
```

```
Query Match      35.2%; Score 1178; DB 2; Length 480;  
Best Local Similarity 50.4%; Pred. No. 6,36-63;  
Matches 242; Conservative 94; Mismatches 130; Indels 14; Gaps 7;
```

```
OY 1 MDHKDLAQRIRLDGGEDNIVAAAHCASTRLELYKDTQDVROSLLDDPDRLGTFEEKGGM 60  
   :::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB 1 MNVYQSADIEDINALCGSEENLDMAHCHTRLRKYLVNDESLVBEHLNMNDVKGFSTFGCG 60  
OY 61 FOIIPEGVDVHFKELEDATSKDIAVTEQLKDVANNANVFSSRAVKVLADI FVPLPI 120  
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB 61 YOIIIGSGTVKVVSLEKLTKGE-ASTTSEVKQSAKNMNPLORFVYMISDI FVPIIPA 119  
OY 121 LVGGGLMAINNVAODL-FGPSIVMEPQISGVAEMLNLMSAPAPLPVVGFPAT 179  
   :|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB 120 IVAGGLMGLNNTITAKDLFFSGSLIDVYSQFGLMEMINVPANAFTLLPIILGISAA 179  
OY 180 KRFGNEFLGAGIGMAWFPPTLVNGDYVAATMTAGE-MPMWSLFGLDVAQAQGYGTLPV 238  
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB 180 KRFGNPFLGAIVLMIIYHPELMSAYDPKPKAVEAGKAIPYWDVGHINQYGQGQYLPM 239  
OY 239 LVSWMILATITEKFHLKXLMGADPLITPVLTLLTLTGFLTAIGPARRWGDLLAHGLQG 298  
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB 240 LVAAYIILASIEKGRKRYIPVLDLMLRPPLSIFITFALETFSVPIITRQLDLYMLSDGLTW 299  
OY 299 LYDGEGPVGGLFGLVNVSPITYTTLHOSFPRIELELF----NOGSSEFI FATASMANIAOG 354  
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB 300 LYREGAMIGLIIFGLYVAPIVTTMHNSFLAVETLLADATKTGSGSFPIPATMSVNAOG 359  
OY 355 AACLAIVFPFLAK-SKKLKGLAGASGVSAVLGITTEPAITGVNIKLMPPIYIGTGTAIGCAL 413  
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB 360 GAALIAAFPIIKONKCLKGVASAADISALLGITTEPAMEGVNIKLYPFI GALVGGIGISAY 419  
OY 414 IALFDIKVALGAGFGVSVSIDA-PDMWFVLCVAVTTFVIAGAIAVGLYVVRNRGS 471  
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB 420 IAFKRVKIALMGTGLGPGTSINVVHAGMWLHYFGMTISFTI-----AITVLLLSKRKAN 475
```

Search completed: March 7, 2005, 22:14:08
Job time : 183 secs

Search completed: March 7, 2005, 22:14:08
Job time : 183 secs

Blue Ridge Blank (uspto)